

UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística e Ciência da Computação

- IMECC -

**ESTIMAÇÃO DO
PONTO DE MUDANÇA DE MÉDIA EM SEQUÊNCIAS DE VARIÁVEIS ALEATÓRIAS**

Elaine Borghi

Orientador : Mauro Sérgio de Freitas Marques

Campinas - São Paulo

1992

ESTIMAÇÃO DO PONTO DE MUDANÇA DE MÉDIA EM SEQUÊNCIAS DE VARIÁVEIS ALEATÓRIAS

Este exemplar corresponde à redação final da dissertação de Mestrado devidamente corrigida e defendida pela Sr^a Elaine Borghi, e aprovada pela comissão julgadora.

Campinas, 17 de julho de 1992.



Prof. Dr. Mauro Sérgio de Freitas Marques
Orientador

Dissertação de Mestrado apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, para obtenção do grau de Mestre em Estatística.

Aos meus pais e irmãos
Dirceu, Vanda, Solange e Dirceu
e à Jane

AGRADECIMENTOS

Ao meu querido professor e orientador Mauro, pela sua demonstração constante de dedicação e paciência e pela segurança transmitida nos momentos difíceis.

À minha família, pela paciência e apoio.

Aos meus amigos, pelo apoio.

Ao Prof. Dr. Oscar Bustos, pela ajuda na parte de simulação.

À professora Maria Eliza Fini, pelo incentivo e pela experiência transmitida.

Àqueles professores do Departamento de Estatística do IMECC que, de alguma forma, colaboraram.

A todos meus colegas do Departamento de Planejamento e Produção Agropecuária da FEAGRI, pelo total apoio, sem o qual este trabalho não seria possível.

À Faculdade de Engenharia Agrícola em geral, pelo incentivo.

Finalmente, meus agradecimentos aos professores Hotta e Dalton, membros da banca de defesa, pelo apoio e grande ajuda com as sugestões dadas.

RESUMO

Consideremos X_1, X_2, \dots, X_n uma sequência de variáveis aleatórias na ordem em que são observadas, sendo $X_1, \dots, X_{[n\theta]}$ têm uma distribuição F e $X_{[n\theta]+1}, \dots, X_n$ têm distribuição G , onde $F \neq G$, $[n\theta]$ é o tempo de mudança com $[y]$ denotando a parte inteira de y e $\theta \in [0,1]$ sendo nosso parâmetro desconhecido.

Este trabalho trata do problema de estimar θ quando a mudança ocorrida está relacionada com o primeiro momento da distribuição, isto é, ocorre mudança na média da distribuição.

Foram considerados os casos em que temos as variáveis aleatórias independentes e foi feita a aplicação do estimador quando temos um processo de ramificação.

SUMÁRIO

Capítulo 1. Introdução	1
Capítulo 2. Estimação do ponto de mudança em uma sequência de variáveis aleatórias independentes	15
2.1. Introdução	15
2.2. Estimação não paramétrica do ponto de mudança da média da distribuição	16
2.3. Consistência do estimador de θ	31
2.4. Aplicação do estimador nos dados do Rio Nilo	53
2.5. Proposta de um segundo estimador	56
2.6. Estudos de simulação	57
2.7. Conclusões	59
Capítulo 3. Aplicação do estimador em processos de ramificação ..	61
3.1. Introdução	61
3.2. Estimação não-paramétrica do ponto de mudança de distribuição em processos de ramificação	63
3.3. Estudo das propriedades do estimador	71
3.4. Estudos de simulação	77

Apêndice A. Tabelas	79
Apêndice B. Gráficos	85
Apêndice C. Aspectos de simulação	91
Apêndice D. Programas	92
Bibliografia	103

CAPÍTULO 1.

INTRODUÇÃO.

É um fato comum ocorrerem alterações durante a observação de um processo aleatório que podem causar mudança na distribuição das variáveis que estão sendo observadas. Estas alterações podem ser causadas por fatores desconhecidos, não observáveis, de forma a não podermos detectar em que momento a mudança ocorreu explicitamente.

Quando o ponto de mudança é conhecido, podemos considerar este problema como uma simples comparação entre tratamentos, mas quando este é desconhecido, o problema torna-se um pouco mais complexo.

Um exemplo prático de aplicação deste problema foi citado em Cobb (1978). Foi registrado o volume anual de descarga do Rio Nilo de 1871 a 1970 e há evidências meteorológicas que ocorreu uma mudança no ano 1898. Assume-se que estas variáveis são independentes com comum variância e normalmente distribuídas em toda a série.

Dada a importância, em algumas áreas tais como biomédica, agronômica ou industrial, do conhecimento dos tempos de mudança de distribuição em determinado processo, alguns autores como

Hinkley (1970), Carlstein (1988) e Perez-Abreu (1989), vêm estudando para certas situações particulares métodos de estimação deste parâmetro, que é o tempo de mudança de distribuição.

Vejamos agora, quando o tempo de mudança é desconhecido, a formulação do problema.

Formulação do problema :

Consideremos X_1, X_2, \dots, X_n uma sequência de variáveis aleatórias na ordem em que são observadas, sendo que $X_1, \dots, X_{[n\theta]}$ têm uma distribuição F e $X_{[n\theta]+1}, \dots, X_n$ têm distribuição G , onde $F \neq G$, $[n\theta]$ é o tempo de mudança ($[y]$ denotando a parte inteira de y) e $\theta \in [0,1]$ sendo nosso parâmetro desconhecido.

Nosso objetivo, neste trabalho, é tratar deste problema quando a mudança ocorrida está relacionada com o primeiro momento da distribuição, isto é, ocorre mudança na média da distribuição.

Ainda, nesta introdução, revisamos estudos feitos por alguns autores nos quais se baseou nossa proposta.

No segundo capítulo apresentamos uma proposta de estimação para o tempo de mudança de média em uma sequência de variáveis aleatórias independentes, mostrando algumas propriedades do estimador. Além disso, aplicaremos o estimador proposto para o exemplo do Rio Nilo, citado anteriormente. Estudos de simulação para verificarmos o comportamento do estimador são apresentados,

fechando este capítulo.

No terceiro capítulo estudaremos o problema no caso de cadeias de Markov. Em particular, estudamos o problema para o caso de um tipo especial de cadeia de Markov, o processo de ramificação, usando o estimador proposto no capítulo 2.

Revisão Bibliográfica

Hinkley (1970) considera uma sequência de variáveis aleatórias independentes e contínuas X_1^n, \dots, X_n^n com função densidade de probabilidade $f(x, \mu_0)$ para X_i^n ($i = 1, \dots, [n\theta]$) e $f(x, \mu_1)$ para X_i^n ($i = [n\theta]+1, \dots, n$), $\theta \in [0,1]$, onde

$$\mu_0 = E(X_i^n), \quad i = 1, \dots, [n\theta],$$

$$\mu_1 = E(X_i^n), \quad i = [n\theta]+1, \dots, n$$

$$\text{e } \theta \in [0,1],$$

ou seja, a mudança de distribuição ocorre apenas no primeiro momento.

Hinkley propõe estimar o parâmetro θ usando o método de máxima verossimilhança. O logaritmo da função de verossimilhança para esta situação é

$$L_n(x_1^n, \mu_0, \mu_1, \lambda) = \sum_{i=1}^{[n\lambda]} \log f(x_i^n, \mu_0) + \sum_{i=[n\lambda]+1}^n \log f(x_i^n, \mu_1)$$

O estimador de máxima verossimilhança para θ , denotado por $\hat{\lambda}_n$ é o valor de λ que maximiza a função $L_n(\tilde{x}_n, \mu_0, \mu_1, \lambda)$, isto é,

$$L_n(\tilde{x}_n, \mu_0, \mu_1, \hat{\lambda}_n) = \max_{\lambda \in [0,1]} L_n(\tilde{x}_n, \mu_0, \mu_1, \lambda)$$

Dado que a expressão para este estimador não é explícita, as propriedades do estimador são estudadas através da distribuição assintótica e de variáveis associadas a passeios aleatórios definidos convenientemente. Assumindo μ_0 e μ_1 conhecidos e $[n\theta]$ e $n-[n\theta]$ suficientemente grandes, Hinkley define

$$U_j^n = \log f(x_j^n, \mu_0) - \log f(x_j^n, \mu_1).$$

Então a função log verossimilhança pôde ser escrita como

$$L_n(\tilde{x}_n, \mu_0, \mu_1, \lambda) = \sum_{j=1}^{[n\lambda]} U_j^n + \sum_{j=1}^n \log f(x_j^n, \mu_1).$$

Assim sendo, o problema de se maximizar $L_n(\tilde{x}_n, \mu_0, \mu_1, \lambda)$ restringe-se em maximizar a função

$$V_{[n\lambda]}^n = \sum_{j=1}^{[n\lambda]} U_j^n, \quad \lambda \in [0,1].$$

Para maximizar esta função, o autor utiliza o fato que

$V_{[n\lambda]}^n - V_{[n\theta]}^n$ define dois passeios aleatórios independentes

$$W = \left\{ 0, - \sum_{j=0}^{[n\lambda]} U_{[n\theta]-j}^n \quad ([n\lambda] = 0, 1, \dots, [n\theta]-1) \right\}$$

$$W' = \left\{ 0, \sum_{j=0}^{[n\lambda]} U_{[n\theta]+1+j}^n \quad ([n\lambda] = 0, 1, \dots, n-[n\theta]-2) \right\},$$

onde W representa a log verossimilhança para inteiros menores que $[n\theta]$ e W' para os inteiros maiores que $[n\theta]$, relativas a log verossimilhança do próprio $[n\theta]$.

Para determinar $\hat{\lambda}_n$, deve-se encontrar λ tal que maximize os dois passeios aleatórios.

A partir disto, o autor chega à distribuição assintótica de $[n\hat{\lambda}_n] - [n\theta]$ para a verificação das propriedades assintóticas de $\hat{\lambda}_n$.

Uma propriedade verificada de $\hat{\lambda}_n$ é a não consistência, isto é, o estimador não converge em probabilidade para o valor real do parâmetro, já que quando n tende para infinito, todos os momentos assintóticos existentes de $[n\hat{\lambda}_n] - [n\theta]$ têm limites não nulos.

Resultados são derivados para o caso particular da distribuição normal. No trabalho citado, o autor deixou claro que estes resultados são aplicáveis a distribuições não normais, inclusive podendo ser generalizados para incluir distribuições discretas.

Esta generalização é apresentada em Hinkley & Hinkley (1970) que aplica estes resultados para uma sequência de variáveis com distribuição Bernoulli.

Especificamente, são consideradas sequências X_1^n, \dots, X_n^n onde X_i^n 's são variáveis aleatórias independentes com distribuição de probabilidade

$$P(X_i^n = 1) = \begin{cases} \mu_0 & , \quad i = 1, \dots, [n\theta] \\ \mu_1 & , \quad i = [n\theta]+1, \dots, n \end{cases}$$

e $P(X_i^n = 0) = 1 - P(X_i^n = 1)$ com μ_0 e μ_1 conhecidos e θ

desconhecido. Como no caso da normal, vemos que aquilo que está sendo considerado é apenas uma mudança na média da distribuição.

Neste trabalho, alguns pontos negativos sobre o estimador $\hat{\lambda}_n$ são destacados, como por exemplo, sua ineficiência como uma estatística para teste, reflexo desta não ser suficiente, o que é uma característica frequente em problemas de pontos de mudança, ainda esclarecem os autores. Também resultados indicam que, para pequenas mudanças no parâmetro binomial, existe um apreciável vício para $\hat{\lambda}_n$.

Carlstein (1988) trata o problema de ponto de mudança de distribuição utilizando funções de distribuições empíricas que não requerem qualquer conhecimento sobre as distribuições das variáveis aleatórias da sequência antes ou depois de ocorrer a mudança de

distribuição, isto é, um estudo não paramétrico.

Para x_1^n, \dots, x_n^n uma amostra de variáveis aleatórias independentes, tem-se que

$X_1^n, \dots, X_{[n\theta]}^n$ são i.i.d. com função de distribuição F

e $X_{[n\theta]+1}^n, \dots, X_n^n$ são i.i.d. com função de distribuição G .

A única suposição em Carlstein (1988) é que o conjunto

$$\Lambda = \{ x \in \mathbb{R} : | F(x) - G(x) | > 0 \}$$

satisfaça

$$\int_{\Lambda} dF(x) > 0 \quad \text{ou} \quad \int_{\Lambda} dG(x) > 0.$$

Sendo assim, pode-se considerar F e G contínuas, discretas ou mistas.

Carlstein assume, hipoteticamente, que o parâmetro θ pertença a um conjunto $T_n = \{ i/n : 1 \leq i \leq n-1 \}$. A partir disso, a função distribuição empírica pré- λ da sequência de variáveis aleatórias i.i.d., denotada por ${}_{\lambda}h^n$ é dada por

$${}_{\lambda}h^n(x) = \sum_{i=1}^{[n\lambda]} 1(X_i^n \leq x) / [n\lambda]$$

e, a pós- λ , denotada por h_{λ}^n , dada por

$$h_{\lambda}^n(x) = \sum_{i=(n\lambda)+1}^n 1(X_i^n \leq x) / (n - [n\lambda]),$$

onde $1(\cdot)$ é a função indicadora.

Tem-se que $h_{\lambda}^n(x)$ e $h_{\lambda}^n(x)$ estimam, de maneira não viciada, as distribuições mistas

$$h_{\lambda}(x) = 1(\lambda \leq \theta) F(x) + 1(\lambda > \theta) (\theta F(x) + (\lambda - \theta) G(x)) / \lambda$$

e

$$h_{\lambda}(x) = 1(\lambda \leq \theta) ((\theta - \lambda) F(x) + (1 - \theta) G(x)) / (1 - \lambda) + \\ + 1(\lambda > \theta) G(x).$$

respectivamente, isto é,

$$E(h_{\lambda}^n(x)) = h_{\lambda}(x) \quad \text{e} \quad E(h_{\lambda}^n(x)) = h_{\lambda}(x).$$

Uma distância entre estas distribuições mistas desconhecidas, $h_{\lambda}(x)$ e $h_{\lambda}(x)$, pode ser medida através das quantidades

$$\begin{aligned} \delta_{ni}^{\lambda} &= |h_{\lambda}^n(X_i^n) - h_{\lambda}(X_i^n)| \\ &= 1(\lambda \leq \theta) \left[\frac{(1 - \lambda - \theta + \lambda)}{(1 - \lambda)} F(X_i^n) - \frac{(1 - \theta)}{(1 - \lambda)} G(X_i^n) \right] + \\ &\quad + 1(\lambda > \theta) \left[\frac{\theta}{\lambda} F(X_i^n) + \frac{(\lambda - \theta - \lambda)}{\lambda} G(X_i^n) \right] = \\ &= 1(\lambda \leq \theta) \frac{(1 - \theta)}{(1 - \lambda)} (F(X_i^n) - G(X_i^n)) + \\ &\quad + 1(\lambda > \theta) \frac{\theta}{\lambda} (F(X_i^n) - G(X_i^n)) \end{aligned}$$

$$\begin{aligned}
&= \left[1(\lambda \leq \theta) \frac{(1-\theta)}{(1-\lambda)} + 1(\lambda > \theta) \frac{\theta}{\lambda} \right] (F(X_i^n) - G(X_i^n)) \\
&= \left[1(\lambda \leq \theta) \frac{(1-\theta)}{(1-\lambda)} + 1(\lambda > \theta) \frac{\theta}{\lambda} \right] \delta_{ni}^{\theta}, \quad 1 \leq i \leq n.
\end{aligned}$$

Combinando estas n diferenças através de uma função norma homogênea

$$S_n: \mathbb{R}^n \longrightarrow \mathbb{R}$$

(i.e., $S_n(cy_1, \dots, cy_n) = cS_n(y_1, \dots, y_n)$, sempre que $c \geq 0$ e $y_i \geq 0$ ($\forall i$)). Carlstein define

$$\begin{aligned}
\Delta_n(\lambda) &= \lambda^{1/2} (1-\lambda)^{1/2} S_n(\delta_{n1}^{\lambda}, \dots, \delta_{nn}^{\lambda}) \\
&= \rho(\lambda) S_n(\delta_{n1}^{\theta}, \dots, \delta_{nn}^{\theta}),
\end{aligned}$$

onde

$$\begin{aligned}
\rho(\lambda) &= 1(\lambda \leq \theta) (1-\theta) \lambda^{1/2} / (1-\lambda)^{1/2} + \\
&\quad + 1(\lambda > \theta) \theta (1-\lambda)^{1/2} / \lambda^{1/2}.
\end{aligned}$$

Pode-se notar que $\Delta_n(\lambda)$ atinge seu máximo quando $\rho(\lambda)$ é maximizada. Este máximo é único e ocorre em $\lambda = \theta$. Baseado neste fato, Carlstein propõe como um estimador razoável para θ o valor $\lambda \in [0,1]$ que maximiza o estimador não viciado de $\Delta_n(\lambda)$, que é a função amostral dada por

$$D_n(\lambda) = \lambda^{1/2} (1-\lambda)^{1/2} S_n(d_{n1}^\lambda, \dots, d_{nn}^\lambda)$$

onde

$$d_{ni}^\lambda = | \lambda h^n OX_i^n - h_\lambda^n OX_i^n |, \quad 1 \leq i \leq n,$$

correspondente a $\Delta_n(\lambda)$.

Denotando-se por $\hat{\lambda}_n$ este estimador, tem-se $\hat{\lambda}_n \in [0,1]$ tal que

$$D_n(\hat{\lambda}_n) = \max_{\lambda \in [0,1]} D_n(\lambda).$$

Carlstein exige ainda que certas propriedades como :

- (i) (Simetria) S_n é simétrica em seus n argumentos,
- (ii) (Homogeneidade) $S_n(cy_n) = c S_n(y_n)$ sempre que $c \geq 0$ e $y_n \geq 0_n$,
- (iii) (Desigualdade triangular) $S_n(y_n^{(1)} + y_n^{(2)}) \leq S_n(y_n^{(1)}) + S_n(y_n^{(2)})$
sempre que $y_n^{(1)} \geq 0_n$ e $y_n^{(2)} \geq 0_n$,
- (iv) (Identidade) $S_n(1_n) = 1$,
- (v) (Monotonicidade) $S_n(y_n^{(1)}) \geq S_n(y_n^{(2)})$ sempre que $y_n^{(1)} \geq y_n^{(2)} \geq 0_n$,
- (vi) (Média dominância) $S_n(y_n) \geq \sum_{i=1}^n y_i / n$ sempre que $y_n \geq 0_n$,

devem ser impostas a norma S_n . Neste caso, a função S_n é dita ser média dominante.

O autor mostra que, assumindo que S_n seja média-dominante, $\hat{\lambda}_n$ é invariante sob transformações estritamente monótonas dos dados, desde que não exista coincidência entre as observações.

Alguns exemplos de normas média-dominantes são dados, como

$$S_n^{(1)}(\tilde{y}^n) = \sum_{i=1}^n y_i^n / n$$

$$S_n^{(2)}(\tilde{y}^n) = \left[\sum_{i=1}^n \frac{y_i^n^2}{n} \right]^{1/2}$$

$$S_n^{(3)}(\tilde{y}^n) = \sup_{1 \leq i \leq n} y_i^n.$$

Além de fazer aplicações, como por exemplo, nos dados do Nilo (Cobb (1978)), é feito também um estudo de simulação, utilizando duas distribuições F e G, ambas simétricas com médias e variâncias similares, tornando-se difícil obter um estimador que detecte a mudança de distribuição. O estimador demonstrou um bom comportamento nos casos estudados e o autor provou a sua consistência.

Outro artigo neste tópico é o de Perez-Abreu (1989), que considera uma sequência de variáveis aleatórias discretas independentes assumindo valores inteiros não negativos e propõe o uso da função geratriz de probabilidades empírica, ao invés da distribuição empírica, para detectar a mudança ocorrida na distribuição num determinado ponto.

Para $X_1^n, \dots, X_{[n\theta]}^n$ independentes com distribuição discreta nos números não negativos comum F e $X_{[n\theta]+1}^n, \dots, X_n^n$ independentes com distribuição comum G, também discretas nos inteiros não negativos, com $F \neq G$, onde $\theta \in (0,1)$, Perez-Abreu

define a função geratriz de probabilidades empírica pré e pós respectivamente por

$${}_{\lambda}\hat{\phi}^n(t) = \frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} t^{X_i^n} \quad t \in [0,1]$$

e

$$\hat{\phi}_{\lambda}^n(t) = \frac{1}{n - [n\lambda]} \sum_{i=[n\lambda]+1}^n t^{X_i^n} \quad t \in [0,1].$$

Tem-se que ${}_{\lambda}\hat{\phi}^n$ é um estimador não viciado para a função

$${}_{\lambda}\phi(t) = \lambda^{-1} \left[\theta \phi_F(t) + (\lambda - \theta) \phi_G(t) \right] 1(\lambda > \theta) + \phi_F(t) 1(\lambda \leq \theta)$$

e $\hat{\phi}_{\lambda}^n$ é um estimador não viciado para a função

$$\begin{aligned} \phi_{\lambda}(t) &= (1-\lambda)^{-1} \left[(\theta - \lambda) \phi_F(t) + (1 - \theta) \phi_G(t) \right] 1(\lambda \leq \theta) \\ &+ \phi_G(t) 1(\lambda > \theta), \end{aligned}$$

isto é,

$$E({}_{\lambda}\hat{\phi}^n(t)) = {}_{\lambda}\phi(t) \quad \text{e} \quad E(\hat{\phi}_{\lambda}^n(t)) = \phi_{\lambda}(t),$$

onde

$$\phi_F(t) = \int t^x dF(x) \quad \text{e} \quad \phi_G(t) = \int t^x dG(x)$$

são as funções geratrizes de probabilidades, referentes às distribuições F e G, respectivamente.

Para $0 \leq \lambda \leq 1$, tem-se que

$$\lambda \phi(t) - \phi_\lambda(t) = \left[\phi_F(t) - \phi_\theta(t) \right] \cdot \left\{ \frac{(1-\theta)}{(1-\lambda)} 1(\lambda \leq \theta) + \frac{\theta}{\lambda} 1(\lambda > \theta) \right\}$$

é uma função que atinge seu valor máximo único quando $\lambda = \theta$.

Considerando este fato, Pérez-Abreu propõe como estimador de θ o valor de λ tal que maximize uma distância entre os estimadores não viciados de $\lambda \phi$ e ϕ_λ .

A distância considerada entre $\lambda \hat{\phi}^n$ e $\hat{\phi}_\lambda^n$ é definida pela função.

$$d_n(\lambda) = \lambda(1-\lambda) \int_0^1 \left[\lambda \hat{\phi}^n(t) - \hat{\phi}_\lambda^n(t) \right]^2 t^2 dt.$$

O fator $\lambda(1-\lambda)$ é introduzido para compensar o fato de que quando λ é pequeno $\lambda \hat{\phi}^n$ não é um bom estimador de $\lambda \phi$ assim como quando λ é grande $\hat{\phi}_\lambda^n$ não é um bom estimador para ϕ_λ . Além disso, se introduz o fator t^2 para diminuir o peso das observações $X_i^n = 0$.

Substituindo-se $\lambda \hat{\phi}^n(t)$ e $\hat{\phi}_\lambda^n(t)$, a distância dada torna-se

$$d_n(\lambda) = \lambda(1-\lambda) \left\{ \frac{1}{(n\lambda)^2} \sum_{i=1}^{n\lambda} \sum_{j=1}^{n\lambda} \frac{1}{X_i^n + X_j^n + 3} + \right. \\ - \frac{2}{n^2 \lambda(1-\lambda)} \sum_{i=1}^{n\lambda} \sum_{j=n\lambda+1}^n \frac{1}{X_i^n + X_j^n + 3} + \\ \left. + \frac{1}{n^2(1-\lambda)^2} \sum_{j=n\lambda+1}^n \sum_{i=n\lambda+1}^n \frac{1}{X_i^n + X_j^n + 3} \right\}.$$

Assim, o estimador proposto para θ é $\hat{\lambda}_n \in [0,1]$ tal que

$$d_n(\hat{\lambda}_n) = \max_{\lambda \in [0,1]} d_n(\lambda).$$

O autor mostra que $\hat{\lambda}_n$ é consistente em probabilidade para θ com razão de convergência exponencial.

É feito também um estudo de simulação Monte Carlo para comparar o estimador proposto por Perez-Abreu com o proposto em Carlstein (1988), utilizando distribuições como Binomial, Geométrica, Poisson e Binomial Negativa com médias diferentes ou similares, bem como com variâncias diferentes ou similares para diferentes tamanhos de amostra e tempos de mudanças θ .

Como conclusões, chega-se que quando o parâmetro de mudança se aproxima dos extremos do intervalo $[0,1]$, o estimador proposto por Perez-Abreu se comporta melhor que o proposto por Carlstein e o contrário ocorre quando este se aproxima de 0.5.

CAPÍTULO 2.

ESTIMAÇÃO DO PONTO DE MUDANÇA DE MÉDIA EM UMA SEQUÊNCIA DE VARIÁVEIS ALEATÓRIAS INDEPENDENTES.

2.1. Introdução.

Como vimos no capítulo anterior, os estimadores propostos por Carlstein (1988) e Perez-Abreu (1989) para o tempo de mudança de distribuição em sequências de variáveis aleatórias independentes, são consistentes e a convergência ocorre em razão exponencial, enquanto que o estimador de máxima verossimilhança estudado por Hinkley (1970) é visto ser não consistente.

Nosso objetivo neste capítulo é propor estimadores para o tempo de mudança na média da distribuição como nos casos considerados em Hinkley (1970) e em Hinkley & Hinkley (1970), porém através de um método não paramétrico, isto é, sem o conhecimento das distribuições envolvidas. Estudaremos algumas propriedades e compararemos estes estimadores com o de máxima verossimilhança estudado por Hinkley.

Considerando que os estudos desenvolvidos em Hinkley (1970) e Hinkley & Hinkley (1970) se restringiram aos casos das distribuições normal e binomial, a título de comparação entre os

estimadores só serão considerados estas duas situações. Além disso, aplicamos nosso estimador usando outras distribuições onde ocorre mudança de média, que foram consideradas em Perez-Abreu (1989).

2.2. Estimação não paramétrica do ponto de mudança da média da distribuição.

Seguindo as propostas de Carlstein (1988) e Perez-Abreu (1989), vamos considerar, para cada possível tempo $[n\lambda]$, $\lambda \in [0,1]$, na sequência de variáveis aleatórias, estatísticas que estimem de maneira apropriada a característica que desejamos detectar a mudança ocorrida. No caso aqui estudado, esta característica é a média. O problema de estimação do tempo de mudança de distribuição $[n\theta]$ pode ser estudado através da comparação dos estimadores das médias estimados nos tempos anterior e posterior a cada possível tempo de mudança deste parâmetro.

Assim sendo, consideremos uma sequência de variáveis aleatórias independentes X_1^n, \dots, X_n^n , na ordem em que foram observadas e $\theta \in (0,1)$ fixo tal que :

$X_1^n, \dots, X_{[n\theta]}^n$ são i.i.d. com função distribuição F

e

$X_{[n\theta]+1}^n, \dots, X_n^n$ são i.i.d. com função distribuição G ,

onde

$$E (X_i^n) = \begin{cases} \mu_0 & , \quad i = 1, \dots, [n\theta] \\ \mu_1 & , \quad i = [n\theta]+1, \dots, n \end{cases}$$

com

$$\mu_0 \neq \mu_1.$$

Vamos assumir também sempre que F e G têm segundos momentos finitos.

Definição 2.2.1. Sejam X_1^n, \dots, X_n^n variáveis aleatórias satisfazendo as condições acima e $\lambda \in (0,1)$. A média amostral pré- λ é a estatística dada por

$$\bar{X}_\lambda^n = \frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} X_i^n$$

e a média amostral pós- λ , a estatística dada por

$$\bar{X}_\lambda^n = \frac{1}{n-[n\lambda]} \sum_{i=[n\lambda]+1}^n X_i^n$$

Proposição 2.2.1. Sejam X_1^n, \dots, X_n^n e $\lambda \in (0,1)$ satisfazendo as condições acima e a função

$$d_\lambda^n = \bar{X}_\lambda^n - \bar{X}_\lambda^n.$$

Então d_λ^n é um estimador não viciado para a função

$$f_{\lambda} = \left\{ \mu_0 - \frac{1}{n - [n\lambda]} \left[([n\theta] - [n\lambda]) \mu_0 + \left(n - [n\theta] \right) \mu_1 \right] \right\} \quad 1(\lambda \leq \theta)$$

$$+ \left\{ \frac{1}{[n\lambda]} \left[[n\theta] \mu_0 + \left([n\lambda] - [n\theta] \right) \mu_1 \right] - \mu_1 \right\} \quad 1(\lambda > \theta).$$

Prova : Dado que a informação que temos sobre a distribuição das variáveis aleatórias da sequência é em função do parâmetro θ , devemos considerar dois casos:

Caso 1: $\lambda \leq \theta$

$$E \left(\bar{X}_{\lambda}^n \right) = E \left[\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} X_i^n \right] = \frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} E(X_i^n) = \frac{1}{[n\lambda]} [n\lambda] \mu_0 = \mu_0.$$

$$E \left(\bar{X}_{\lambda}^n \right) = E \left[\frac{1}{n - [n\lambda]} \sum_{i=[n\lambda]+1}^n X_i^n \right] = E \left[\frac{1}{n - [n\lambda]} \left(\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^n X_i^n \right) \right] =$$

$$= \frac{1}{n - [n\lambda]} \left(\sum_{i=[n\lambda]+1}^{[n\theta]} E(X_i^n) + \sum_{i=[n\theta]+1}^n E(X_i^n) \right) =$$

$$= \frac{1}{n - [n\lambda]} \left[\left([n\theta] - [n\lambda] \right) \mu_0 + \left(n - [n\theta] \right) \mu_1 \right].$$

Temos então que, no caso $\lambda \leq \theta$,

$$E \langle d_{\lambda}^n \rangle = \mu_0 - \frac{1}{n - [n\lambda]} \left[\left([n\theta] - [n\lambda] \right) \mu_0 + \left(n - [n\theta] \right) \mu_1 \right].$$

Caso 2: $\lambda > \theta$

$$\begin{aligned} E \langle \bar{X}_{\lambda}^n \rangle &= E \left[\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} X_i^n \right] = \frac{1}{[n\lambda]} E \left[\sum_{i=1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^{[n\lambda]} X_i^n \right] = \\ &= \frac{1}{[n\lambda]} E \left[\sum_{i=1}^{[n\theta]} E \langle X_i^n \rangle + \sum_{i=[n\theta]+1}^{[n\lambda]} E \langle X_i^n \rangle \right] = \\ &= \frac{1}{[n\lambda]} \left[[n\theta] \mu_0 + \left([n\lambda] - [n\theta] \right) \mu_1 \right]. \end{aligned}$$

$$\begin{aligned} E \langle \bar{X}_{\lambda}^n \rangle &= E \left[\frac{1}{n - [n\lambda]} \sum_{i=[n\lambda]+1}^n X_i^n \right] = \frac{1}{n - [n\lambda]} \sum_{i=[n\lambda]+1}^n E \langle X_i^n \rangle = \\ &= \frac{1}{n - [n\lambda]} \left(n - [n\lambda] \right) \mu_1 = \\ &= \mu_1. \end{aligned}$$

Segue que, para $\lambda > \theta$,

$$E \langle d_{\lambda}^n \rangle = \frac{1}{[n\lambda]} \left[[n\theta] \mu_0 + \left([n\lambda] - [n\theta] \right) \mu_1 \right] = \mu_1.$$

Combinando caso (1) com caso (2), temos :

$$\begin{aligned}
 E \left(d_{\lambda}^n \right) &= \left\{ \mu_0 - \frac{1}{n - [n\lambda]} \left[([n\theta] - [n\lambda]) \mu_0 + (n - [n\theta]) \mu_1 \right] \right\} \quad 1 < \lambda \leq \theta > \\
 &+ \left\{ \frac{1}{[n\lambda]} \left[[n\theta] \mu_0 + ([n\lambda] - [n\theta]) \mu_1 \right] - \mu_1 \right\} \quad 1 < \lambda > \theta > = \\
 &= f_{\lambda}.
 \end{aligned}$$

Agora, dado que estamos considerando a sequência finita X_1, \dots, X_n para a estimação do tempo de mudança e, que na verdade estamos propondo um estimador para $[n\theta]$, podemos restringir, para cada n , os valores de λ ao conjunto

$$T_n = \left\{ i/n ; i = 1, \dots, n-1 \right\}.$$

Sob esta restrição, f_{λ} é dada por

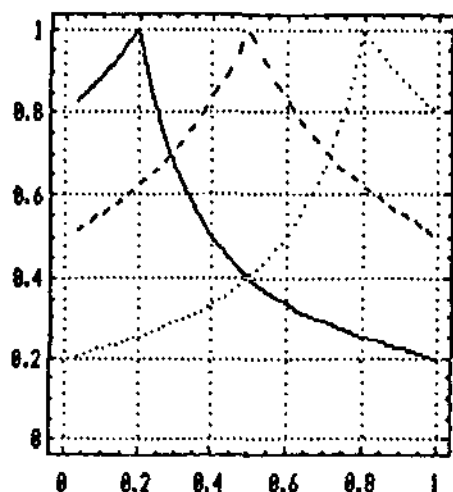
$$\begin{aligned}
 f_{\lambda} &= \left\{ \mu_0 - \frac{(\theta - \lambda)}{(1 - \lambda)} \mu_0 - \frac{(1 - \theta)}{(1 - \lambda)} \mu_1 \right\} \quad 1 < \lambda \leq \theta > + \\
 &+ \left\{ \frac{\theta}{\lambda} \mu_0 + \frac{(\lambda - \theta)}{\lambda} \mu_1 - \mu_1 \right\} \quad 1 < \lambda > \theta > =
 \end{aligned}$$

$$\begin{aligned}
&= \left\{ \frac{(1-\lambda-\theta+\lambda)}{(1-\lambda)} \mu_0 - \frac{(1-\theta)}{(1-\lambda)} \mu_1 \right\} 1(\lambda \leq \theta) + \\
&+ \left\{ \frac{\theta}{\lambda} \mu_0 + \frac{(\lambda-\theta-\lambda)}{\lambda} \mu_1 \right\} 1(\lambda > \theta) = \\
&= \frac{(1-\theta)}{(1-\lambda)} (\mu_0 - \mu_1) 1(\lambda \leq \theta) + \frac{\theta}{\lambda} (\mu_0 - \mu_1) 1(\lambda > \theta) = \\
&= (\mu_0 - \mu_1) \left\{ \frac{(1-\theta)}{(1-\lambda)} 1(\lambda \leq \theta) + \frac{\theta}{\lambda} 1(\lambda > \theta) \right\}.
\end{aligned}$$

Observando os gráficos 2.2.1 e 2.2.2 abaixo, podemos notar que, quando $\mu_0 > \mu_1$ (Gráfico 2.2.1), a função f_λ é crescente no intervalo $(0, \theta]$ e decrescente no intervalo $(\theta, 1)$, atingindo seu valor máximo em $\lambda = \theta$. Ao contrário, quando $\mu_0 < \mu_1$ (Gráfico 2.2.2), a função f_λ é decrescente no intervalo $(0, \theta]$ e crescente no intervalo $(\theta, 1)$, atingindo seu valor mínimo em $\lambda = \theta$.

Gráfico 2.2.1. $\mu_0 \geq \mu_1$

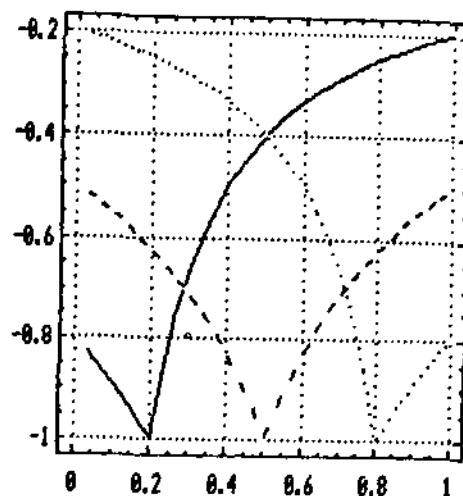
$$f_\lambda / (\mu_0 - \mu_1)$$



λ

Gráfico 2.2.2. $\mu_0 < \mu_1$

$$f_\lambda / (\mu_0 - \mu_1)$$



λ

Legenda :
 — $\theta = 0.2$
 - - - $\theta = 0.5$
 . . . $\theta = 0.8$

Vamos introduzir como notação,

$${}_{\lambda}\mu = E ({}_{\lambda}\bar{X}^n) \quad \text{e} \quad \mu_{\lambda} = E (\bar{X}_{\lambda}^n),$$

isto é,

$$f_{\lambda} = {}_{\lambda}\mu - \mu_{\lambda}.$$

Considerando o fato de que d_{λ}^n estima de maneira não viciada a função f_{λ} , é razoável supormos que um ponto que maximiza a função

amostral $|d_\lambda^n|$ esteja bem próximo do ponto que maximiza $|f_\lambda|$, isto é, este ponto de máximo pode ser considerado um estimador para o ponto que maximiza f_λ . Restringindo-nos a T_n , podemos considerar como um possível estimador de θ o valor $\hat{\lambda}_n \in T_n$ que maximize a distância amostral correspondente dada por

$$D_\lambda^n = |d_\lambda^n| = |\lambda \bar{X}^n - \bar{X}_\lambda^n|$$

Vamos introduzir, também, a seguinte notação

$$\Delta_\lambda = |f_\lambda| = |\lambda\mu - \mu_\lambda|.$$

Um estudo mais detalhado acerca deste estimador passa a ser não trivial, dado que o mesmo não possui uma forma explícita.

Como assumimos que as distribuições têm segundo momento finito, podemos estudar a variância de d_λ^n para cada λ . Esta é dada por

$$\text{Var} (d_\lambda^n) = \text{Var} (\lambda \bar{X}^n - \bar{X}_\lambda^n) = \text{Var} (\lambda \bar{X}^n) + \text{Var} (\bar{X}_\lambda^n),$$

já que X_i e X_j são independentes para $i \neq j$. Nenhuma suposição sobre estas variâncias serem iguais está sendo considerada, portanto elas podem ser distintas. O fato é que estamos interessados somente na média. Vamos denotar por σ_0^2 a variância das variáveis aleatórias antes de ocorrer a mudança e por σ_1^2 , depois da mudança. Como no caso da esperança, devemos considerar duas situações :

Caso 1: $\lambda \leq \theta$.

$$\begin{aligned}
 \text{Var} \left(\bar{X}_\lambda^n \right) &= E \left[\left(\bar{X}_\lambda^n \right)^2 \right] - E^2 \left(\bar{X}_\lambda^n \right) = \\
 &= E \left[\frac{1}{[n\lambda]^2} \sum_{i=1}^{[n\lambda]} X_i^n \sum_{j=1}^{[n\lambda]} X_j^n \right] - E^2 \left[\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} X_i^n \right] = \\
 &= \frac{1}{[n\lambda]^2} \left\{ \sum_{i=1}^{[n\lambda]} \sum_{\substack{j=1 \\ i \neq j}}^{[n\lambda]} E \left(X_i^n X_j^n \right) + \sum_{i=1}^{[n\lambda]} E \left(X_i^n \right)^2 - \left[\sum_{i=1}^{[n\lambda]} E \left(X_i^n \right) \right]^2 \right\} = \\
 &= \frac{1}{[n\lambda]^2} \left\{ \sum_{i=1}^{[n\lambda]} \sum_{\substack{j=1 \\ i \neq j}}^{[n\lambda]} E \left(X_i^n \right) E \left(X_j^n \right) + \sum_{i=1}^{[n\lambda]} E \left(X_i^n \right)^2 - \left[\sum_{i=1}^{[n\lambda]} E \left(X_i^n \right) \right]^2 \right\} = \\
 &= \frac{1}{[n\lambda]^2} \left\{ \left([n\lambda]^2 - [n\lambda] \right) \mu_0^2 + [n\lambda] \left(\sigma_0^2 + \mu_0^2 \right) - [n\lambda]^2 \mu_0^2 \right\} = \\
 &= \frac{\sigma_0^2}{[n\lambda]}.
 \end{aligned}$$

$$\begin{aligned}
 \text{Var} \left(\bar{X}_\lambda^n \right) &= E \left[\left(\bar{X}_\lambda^n \right)^2 \right] - E^2 \left(\bar{X}_\lambda^n \right) = \\
 &= E \left[\frac{1}{(n-[n\lambda])^2} \sum_{i=[n\lambda]+1}^n X_i^n \sum_{j=[n\lambda]+1}^n X_j^n \right] - E^2 \left[\frac{1}{n-[n\lambda]} \sum_{i=[n\lambda]+1}^n X_i^n \right] = \\
 &= \frac{1}{(n-[n\lambda])^2} \left\{ \left[\left(\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^n X_i^n \right) \left(\sum_{j=[n\lambda]+1}^{[n\theta]} X_j^n + \sum_{j=[n\theta]+1}^n X_j^n \right) \right] \right\}
 \end{aligned}$$

$$\begin{aligned}
& - E^2 \left[\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^n X_i^n \right] \Bigg\} = \\
& = \frac{1}{(n-[n\lambda])^2} \left\{ E \left[\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n \sum_{j=[n\lambda]+1}^{[n\theta]} X_j^n \right] + 2 E \left[\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n \sum_{j=[n\theta]+1}^n X_j^n \right] \right. \\
& \quad + E \left[\sum_{i=[n\theta]+1}^n X_i^n \sum_{j=[n\theta]+1}^n X_j^n \right] - E^2 \left[\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^n X_i^n \right] = \\
& = \frac{1}{(n-[n\lambda])^2} \left\{ \sum_{\substack{i=[n\lambda]+1 \\ i \neq j}}^{[n\theta]} \sum_{j=[n\lambda]+1}^{[n\theta]} E \langle X_i^n \rangle E \langle X_j^n \rangle + \sum_{i=[n\lambda]+1}^{[n\theta]} E \langle X_i^n \rangle^2 \right. \\
& \quad + 2 \sum_{i=[n\lambda]+1}^{[n\theta]} \sum_{\substack{j=[n\theta]+1 \\ i \neq j}}^n E \langle X_i^n \rangle E \langle X_j^n \rangle + \sum_{i=[n\theta]+1}^n \sum_{j=[n\theta]+1}^n E \langle X_i^n \rangle E \langle X_j^n \rangle \\
& \quad + \sum_{i=[n\theta]+1}^n E \langle X_i^n \rangle^2 - E^2 \left[\sum_{i=[n\lambda]+1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^n X_i^n \right] = \\
& = \frac{1}{(n-[n\lambda])^2} \left\{ \left[([n\theta] - [n\lambda])^2 - ([n\theta] - [n\lambda]) \right] \mu_0^2 \right. \\
& \quad + ([n\theta] - [n\lambda]) (\sigma_0^2 + \mu_0^2) + 2 ([n\theta] - [n\lambda]) (n - [n\theta]) \mu_0 \mu_1 \\
& \quad + \left[(n - [n\theta])^2 - (n - [n\theta]) \right] \mu_1^2 + (n - [n\theta]) (\sigma_1^2 + \mu_1^2) \\
& \quad - ([n\theta] - [n\lambda])^2 \mu_0^2 - 2 ([n\theta] - [n\lambda]) (n - [n\theta]) \mu_0 \mu_1 \\
& \quad \left. - (n - [n\theta])^2 \mu_1^2 \right\} =
\end{aligned}$$

$$= \frac{1}{(n - [n\lambda])^2} \left\{ \left([n\theta] - [n\lambda] \right) \sigma_1^2 + (n - [n\theta]) \sigma_0^2 \right\}.$$

Caso 2 : $\lambda > \theta$

$$\text{Var} \left(\bar{X}_\lambda^n \right) = E \left[\left(\bar{X}_\lambda^n \right)^2 \right] - E^2 \left(\bar{X}_\lambda^n \right) =$$

$$= E \left[\frac{1}{[n\lambda]^2} \sum_{i=1}^{[n\lambda]} X_i^n \sum_{j=1}^{[n\lambda]} X_j^n \right] - E^2 \left[\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} X_i^n \right] =$$

$$= \frac{1}{[n\lambda]^2} \left\{ E \left[\left(\sum_{i=1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^{[n\lambda]} X_i^n \right) \left(\sum_{j=1}^{[n\theta]} X_j^n + \sum_{j=[n\theta]+1}^{[n\lambda]} X_j^n \right) \right] \right. \\ \left. - E^2 \left(\sum_{i=1}^{[n\theta]} X_i^n + \sum_{i=[n\theta]+1}^{[n\lambda]} X_i^n \right) \right\} =$$

$$= \frac{1}{[n\lambda]^2} \left\{ E \left[\sum_{i=1}^{[n\theta]} X_i^n \sum_{j=1}^{[n\theta]} X_j^n \right] + 2 E \left[\sum_{i=1}^{[n\theta]} X_i^n \sum_{j=[n\theta]+1}^{[n\lambda]} X_j^n \right] \right. \\ \left. + E \left[\sum_{i=[n\theta]+1}^{[n\lambda]} X_i^n \sum_{j=[n\theta]+1}^{[n\lambda]} X_j^n \right] - \left[\sum_{i=1}^{[n\theta]} E \langle X_i^n \rangle + \sum_{i=[n\theta]+1}^{[n\lambda]} E \langle X_i^n \rangle \right]^2 \right\} =$$

$$= \frac{1}{[n\lambda]^2} \left\{ \left([n\theta]^2 - [n\theta] \right) \mu_0^2 + [n\theta] \left(\sigma_0^2 + \mu_0^2 \right) \right.$$

$$\left. + 2 [n\theta] \left([n\lambda] - [n\theta] \right) \mu_0 \mu_1 + \left(\left([n\lambda] - [n\theta] \right)^2 - ([n\lambda] - [n\theta]) \right) \mu_1^2 \right\}$$

$$+ ([n\lambda] - [n\theta]) (\sigma_1^2 + \mu_1^2) - [n\theta]^2 \mu_0^2 - 2 [n\theta] ([n\lambda] - [n\theta]) \mu_0 \mu_1$$

$$- ([n\lambda] - [n\theta])^2 \mu_1^2 \Big\} =$$

$$= \frac{1}{[n\lambda]^2} \left\{ [n\theta] \sigma_0^2 + ([n\lambda] - [n\theta]) \sigma_1^2 \right\}.$$

$$\text{Var} (\bar{X}_\lambda^n) = E \left[(\bar{X}_\lambda^n)^2 \right] - E^2 (\bar{X}_\lambda^n) =$$

$$= E \left[\frac{1}{(n - [n\lambda])^2} \sum_{i=[n\lambda]+1}^n X_i^n \sum_{j=[n\lambda]+1}^n X_j^n \right] - E^2 \left[\frac{1}{n - [n\lambda]} \sum_{i=[n\lambda]+1}^n X_i^n \right] =$$

$$= \frac{1}{(n - [n\lambda])^2} \left\{ \sum_{\substack{i=[n\lambda]+1 \\ i \neq j}}^n \sum_{j=[n\lambda]+1}^n E (X_i^n) E (X_j^n) + \sum_{i=[n\lambda]+1}^n E (X_i^n)^2 - \left[\sum_{i=[n\lambda]+1}^n E (X_i^n) \right]^2 \right\} =$$

$$= \frac{1}{(n - [n\lambda])^2} \left\{ \left[(n - [n\lambda])^2 - (n - [n\lambda]) \right] \mu_1^2 + (n - [n\lambda]) (\sigma_1^2 + \mu_1^2) - (n - [n\lambda])^2 \mu_1^2 \right\} =$$

$$= \frac{\sigma_1^2}{(n - [n\lambda])}.$$

Portanto, combinando as duas situações acima, temos que

$$\text{Var} (d_{\lambda}^n) = \left\{ \frac{\sigma_0^2}{[n\lambda]} + \frac{1}{(n-[n\lambda])^2} \left[([n\theta]-[n\lambda])\sigma_0^2 + (n-[n\theta])\sigma_1^2 \right] \right\} 1(0 \leq \theta) \\ + \left\{ \frac{1}{[n\lambda]^2} \left[[n\theta] \sigma_0^2 + ([n\lambda]-[n\theta])\sigma_1^2 \right] + \frac{\sigma_1^2}{(n-[n\lambda])} \right\} 1(0 > \theta).$$

Como podemos observar, a variância de d_{λ}^n tende para infinito quando λ tende ao valor zero ou um. Portanto, temos um efeito de cauda. Seria razoável considerarmos uma correção para conter esta variância nos extremos.

Uma possível correção pode ser feita fazendo

$$d_{\lambda}^{n*} = \lambda(1-\lambda) d_{\lambda}^n.$$

Neste caso, vejamos o que acontece com a função

$$f_{\lambda}^* = \lambda(1-\lambda) f_{\lambda},$$

que será estimada de maneira não viciada pela função amostral d_{λ}^{n*} .

Observando os gráficos 2.2.3 e 2.2.4 abaixo, podemos notar que ocorre uma linearização por partes na função transformada.

Por outro lado, a variância de d_{λ}^{n*} é dada por

$$\text{Var} (d_{\lambda}^{n*}) = \text{Var} (\lambda(1-\lambda) d_{\lambda}^n) = \left[\lambda(1-\lambda) \right]^2 \text{Var} (d_{\lambda}^n).$$

Então, temos que, quando λ assume os valores zero ou um, a variância de d_{λ}^{n*} é igual a zero.

Gráfico 2.2.3. $\mu_0 \geq \mu_1$

$$f_{\lambda}^*/(\mu_0 - \mu_1)$$

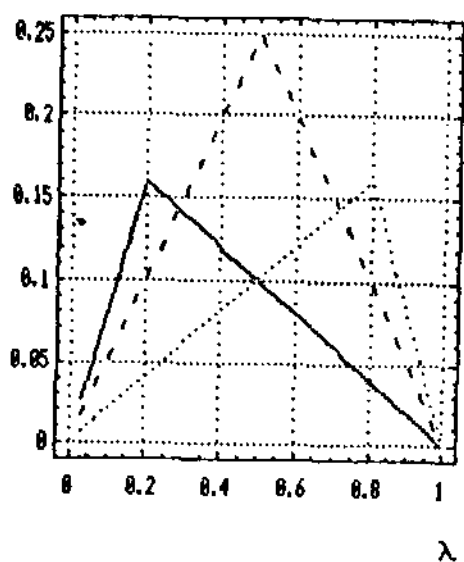
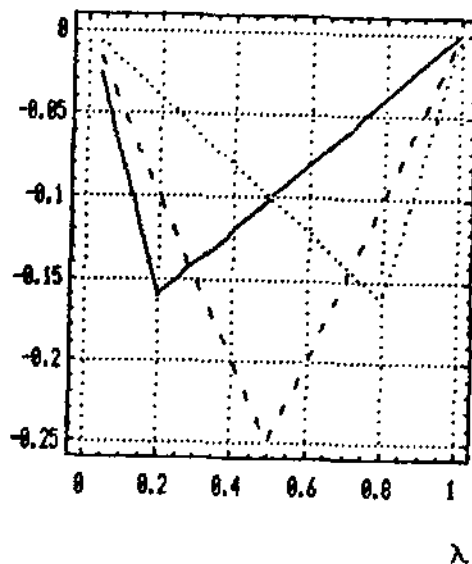


Gráfico 2.2.4. $\mu_0 < \mu_1$

$$f_{\lambda}^*/(\mu_0 - \mu_1)$$



Legenda :
 — $\theta = 0.2$
 - - - $\theta = 0.5$
 . . . $\theta = 0.8$

Relevando estes fatores, passemos, então, a considerar a função amostral dada por

$$D_{\lambda}^{n*} = |d_{\lambda}^{n*}| = \lambda(1-\lambda) |d_{\lambda}^n|.$$

Portanto, passamos a considerar como estimador de θ , $\hat{\lambda}_n$ tal que

$$D_{\hat{\lambda}_n}^{n*} = \max_{\lambda \in T_n} D_{\lambda}^{n*}.$$

Vamos considerar também a função Δ_{λ}^* como sendo

$$\Delta_{\lambda}^* = \lambda(1-\lambda) \Delta_{\lambda}.$$

Proposição 2.2.2. d_{λ}^{n*} é estimador consistente de f_{λ}^* para $\lambda \in [0,1]$.

Prova : Como já vimos anteriormente, d_{λ}^{n*} é estimador não viciado para f_{λ}^* para cada valor fixo $\lambda \in (0,1)$. Sêgue, por Tchebychev e pela expressão da variância de d_{λ}^{n*} , que

$$\begin{aligned} P \left[|d_{\lambda}^{n*} - f_{\lambda}^*| > \varepsilon \right] &\leq \frac{E \left[|d_{\lambda}^{n*} - f_{\lambda}^*|^2 \right]}{\varepsilon^2} = \frac{\text{Var}(d_{\lambda}^{n*})}{\varepsilon^2} = \\ &= \frac{1}{n\varepsilon^2} \left\{ Q(\lambda, \theta) \right\} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

onde $Q(\lambda, \theta) = n \text{Var}(d_{\lambda}^{n*})$. Portanto, d_{λ}^{n*} converge em probabilidade a f_{λ}^* para $\lambda \in [0,1]$. ■

2.3. Consistência do estimador de θ .

Como nosso estimador $\hat{\lambda}_n$ não pode ser escrito de forma explícita, fica difícil verificar suas propriedades.

Uma propriedade importante deste estimador é a sua consistência. Para verificarmos isto, será utilizado argumentos análogos aos que Carlstein e Perez-Abreu. Os resultados que deverão ser utilizados são apresentados a seguir na forma de lemas e da seguinte versão do Teorema de Cramér :

Teorema de Cramér : Seja Y_1, Y_2, \dots , variáveis aleatórias i.i.d. com distribuição H e $E(Y_1) = \nu$. Suponhamos que a função geratriz de momentos

$$M(\theta) = \int e^{\theta x} dH(x)$$

é finita para todo θ . Então $(\forall \epsilon > 0), \exists J(\epsilon) > 0$ tal que

$$\lim_{n \rightarrow \infty} n^{-1} \log P \left[\left| \sum_{i=1}^n \frac{Y_i}{n} - \nu \right| > \epsilon \right] = -J(\epsilon).$$

Referência : Varadhan (1984).

As hipóteses do Teorema de Cramér são satisfeitas para as distribuições usuais como, por exemplo, normal, binomial, exponencial, etc. (Veja, por exemplo, Bahadur (1971)).

Para os próximos lemas, vamos assumir que as sequências de

variáveis consideradas satisfaçam as condições do Teorema de Cramér.

Lema 2.3.1 : Seja $Y_1^{(n)}, \dots, Y_n^{(n)}$ i.i.d. com média μ e r, s, l, m elementos de $\bar{T}_n = T_n \cup \{0, 1\}$ que satisfazem $m \leq s < r \leq l$. Então $(\forall \varepsilon > 0), \exists J(\varepsilon)$ positivo tal que

$$P \left(\sup_{\substack{m, l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| > \varepsilon \right) \leq e^{-n J(\varepsilon)}.$$

Prova : Temos que

$$\left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| = \frac{(r-s)}{(l-m)} \left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right|.$$

Como $(r-s) \leq (l-m)$, segue que

$$\left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| \leq \left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right|.$$

Então

$$\sup_{\substack{m, l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| \leq \left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right|.$$

Por outro lado, temos que

$$\left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right| \leq \sup_{\substack{m,l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right|.$$

Portanto, temos a igualdade :

$$\begin{aligned} \sup_{\substack{m,l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| &= \\ &= \left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right|. \end{aligned} \quad (\text{eq.2.3.1.1})$$

Segue que

$$\begin{aligned} P \left(\sup_{\substack{m,l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(l-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(l-m)} \mu \right| > \varepsilon \right) &= \\ &= P \left(\left| \frac{1}{n(r-s)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \mu \right| > \varepsilon \right) = P_n. \end{aligned} \quad (2.3.1.2)$$

Consideremos a sequência de variáveis Z_1, Z_2, \dots da seguinte forma :

$$Z_i^n = Y_{sn+i}^n, \quad i = 1, \dots, rn - sn$$

e

$$M = rn - sn.$$

Temos que

$$P_n = P \left[\left| \frac{1}{M} \sum_{i=1}^M Z_i^{(n)} - \mu \right| > \varepsilon \right]$$

e, aplicando o Teorema de Cramér, temos

$$\lim_{M \rightarrow \infty} M^{-1} \log P \left[\left| \frac{1}{M} \sum_{i=1}^M Z_i^{(n)} - \mu \right| > \varepsilon \right] = -J(\varepsilon).$$

Segue que

$$P \left[\left| \frac{1}{M} \sum_{i=1}^M Z_i^{(n)} - \mu \right| > \varepsilon \right] \leq e^{-MJ'(\varepsilon)} = e^{-n(r-s)J'(\varepsilon)} = e^{-nJ''(\varepsilon)}$$

e, considerando (2.3.1.2) , completamos a prova do lema. ■

Lema 2.3.2.:

$$P \left\{ \sup_{t \in T_n} | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} \xrightarrow[n \rightarrow \infty]{} 0.$$

Prova : Vamos utilizar a seguinte notação :

$${}_t H^{(n)} = | {}_t \bar{X}^{(n)} - {}_t \mu | \quad \text{e} \quad H_t^{(n)} = | \bar{X}_t^{(n)} - \mu_t |$$

onde, lembrando,

$${}_t \mu = E ({}_t \bar{X}^{(n)}) \quad \text{e} \quad \mu_t = E (\bar{X}_t^{(n)}) ,$$

${}_t \bar{X}^{(n)}$ e $\bar{X}_t^{(n)}$ são as médias amostrais pré-t e pós-t,

respectivamente, e

$$e_t^{(n)} = {}_t H^{(n)} + H_t^{(n)} .$$

Fazendo

$$\begin{aligned} D_t^{(n)} &= | {}_t \bar{X}^{(n)} - \bar{X}_t^{(n)} | = | ({}_t \bar{X}^{(n)} - {}_t \mu) - (\bar{X}_t^{(n)} - \mu_t) - (\mu_t - {}_t \mu) | \leq \\ &\leq | {}_t \bar{X}^{(n)} - {}_t \mu | + | {}_t \mu - \mu_t | + | \mu_t - \bar{X}_t^{(n)} | = \\ &= {}_t H^{(n)} + \Delta_t + H_t^{(n)} . \end{aligned}$$

Segue que

$$D_t^{(n)} - \Delta_t \leq e_t^{(n)} .$$

Agora,

$$\begin{aligned}\Delta_t &= | {}_t\mu - \mu_t | = | \langle {}_t\bar{X}^{(n)} - {}_t\mu \rangle - \langle \bar{X}_t^{(n)} - \mu_t \rangle - \langle {}_t\bar{X}^{(n)} - \bar{X}_t^{(n)} \rangle | \leq \\ &\leq | {}_t\bar{X}^{(n)} - {}_t\mu | + | \bar{X}_t^{(n)} - \mu_t | + | {}_t\bar{X}^{(n)} - \bar{X}_t^{(n)} | = e_t^{(n)} + D_t^{(n)},\end{aligned}$$

isto é,

$$D_t^{(n)} - \Delta_t \geq -e_t^{(n)},$$

o que implica em

$$| D_t^{(n)} - \Delta_t | \leq e_t^{(n)} = {}_tH^{(n)} + H_t^{(n)}.$$

Por outro lado,

$${}_t\mu = \mu_0 \quad 1(t \leq \theta) + \left\{ \frac{\theta}{t} \mu_0 + \frac{(t-\theta)}{t} \mu_1 \right\} \quad 1(t > \theta)$$

Portanto, temos

$$\begin{aligned}{}_tH^{(n)} &= | {}_t\bar{X}^{(n)} - {}_t\mu | \leq 1(t \leq \theta) \left| \frac{1}{nt} \sum_{i=1}^{nt} X_i^{(n)} - \mu_0 \right| + \\ &+ 1(t > \theta) \left[\left| \frac{1}{nt} \sum_{i=1}^{[n\theta]} X_i^{(n)} - \frac{\theta}{t} \mu_0 \right| + \left| \frac{1}{nt} \sum_{i=[n\theta]+1}^{nt} X_i^{(n)} - \frac{(t-\theta)}{t} \mu_1 \right| \right] \leq\end{aligned}$$

$$\begin{aligned}
&\leq 1(t \leq \theta) \left| \frac{1}{nt} \sum_{i=1}^{nt} X_i^{(n)} - \mu_0 \right| + \\
&+ 1(t > \theta) \left[\left| \frac{1}{nt} \sum_{i=1}^{[n\theta]} X_i^{(n)} - \frac{[n\theta]}{nt} \mu_0 \right| + \right. \\
&+ \frac{\left| \frac{[n\theta]}{n} - \theta \right|}{t} \mu_0 + \left| \frac{1}{nt} \sum_{i=[n\theta]+1}^{nt} X_i^{(n)} - \left(1 - \frac{[n\theta]}{nt}\right) \mu_1 \right| + \\
&\left. + \frac{\left| \frac{[n\theta]}{n} - \theta \right|}{t} \mu_1 \right] =
\end{aligned}$$

$$\begin{aligned}
&= 1(t \leq \theta) \left| \frac{1}{nt} \sum_{i=1}^{nt} X_i^{(n)} - \mu_0 \right| + \\
&+ 1(t > \theta) \left[\left| \frac{1}{[n\theta]} \sum_{i=1}^{[n\theta]} X_i^{(n)} - \mu_0 \right| \left(\frac{[n\theta]}{nt} \right) + \right. \\
&+ \left| \frac{1}{nt - [n\theta]} \sum_{i=[n\theta]+1}^{nt} X_i^{(n)} - \mu_1 \right| \left(1 - \frac{[n\theta]}{nt} \right) +
\end{aligned}$$

$$+ \left(\frac{n\theta - [n\theta]}{nt} \right) (\mu_0 + \mu_1) \Bigg] = H_1.$$

Como, quando $t > \theta$, $[n\theta] < nt$ e $0 < n\theta - [n\theta] < 1$ e, utilizando (2.3.1.1), temos então que

$$\begin{aligned} H_1 &\leq 1(t \leq \theta) \sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq t}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{nt} Y_i^{(n)} - \frac{t}{(1-m)} \mu_0 \right| + \\ &+ 1(t > \theta) \left[\sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq [n\theta]/n}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{[n\theta]} Y_i^{(n)} - \frac{([n\theta]/n)}{(1-m)} \mu_0 \right| + \right. \\ &+ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq [n\theta]/n \\ l \geq t}} \left| \frac{1}{n(1-m)} \sum_{i=[n\theta]+1}^{nt} Y_i^{(n)} - \frac{(t-[n\theta]/n)}{(1-m)} \mu_1 \right| + \\ &+ \frac{1}{n} \left(\frac{(\mu_0 + \mu_1)}{t} \right) \Bigg] = {}_t H^{(n)*}. \end{aligned}$$

Temos então que ${}_t H^{(n)*}$ é uma cota superior para ${}_t H^{(n)}$.

Precisamos, agora, de uma cota para $H_t^{(n)}$. Mas,

$$H_t^{(n)} = \left| \bar{X}_t^{(n)} - \mu_t \right|$$

e

$$\mu_t = \left\{ \frac{(\theta-t)}{(1-t)} \mu_0 + \frac{(1-\theta)}{(1-t)} \mu_1 \right\} \mathbf{1}(t \leq \theta) + \mu_1 \mathbf{1}(t > \theta),$$

Portanto, temos

$$\begin{aligned} H_t^{(n)} &= \left| \left[\frac{1}{(n-nt)} \left(\sum_{i=nt+1}^{[n\theta]} X_i^{(n)} + \sum_{i=[n\theta]+1}^n X_i^{(n)} \right) - \left(\frac{\theta-t}{1-t} \right) \mu_0 - \left(\frac{1-\theta}{1-t} \right) \mu_1 \right] \mathbf{1}(t \leq \theta) + \right. \\ &\quad \left. + \left[\frac{1}{(n-nt)} \left(\sum_{i=nt+1}^n X_i^{(n)} - \mu_1 \right) \right] \mathbf{1}(t > \theta) \right| \leq \\ &\leq \mathbf{1}(t \leq \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^{[n\theta]} X_i^{(n)} - \left(\frac{\theta-t}{1-t} \right) \mu_0 \right| + \\ &+ \mathbf{1}(t \leq \theta) \left| \frac{1}{(n-nt)} \sum_{i=[n\theta]+1}^n X_i^{(n)} - \left(\frac{1-\theta}{1-t} \right) \mu_1 \right| + \end{aligned}$$

$$\begin{aligned}
& + 1(t > \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^n X_i^{(n)} - \mu_1 \right| = \\
& = 1(t \leq \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^{[n\theta]} X_i^{(n)} - \frac{([n\theta]-nt)}{n(1-t)} \mu_0 - \frac{(n\theta-[n\theta])}{n(1-t)} \mu_0 \right| + \\
& + 1(t \leq \theta) \left| \frac{1}{(n-nt)} \sum_{i=[n\theta]+1}^n X_i^{(n)} - \frac{(n-[n\theta])}{n(1-t)} \mu_1 - \frac{([n\theta]-n\theta)}{n(1-t)} \mu_1 \right| + \\
& + 1(t > \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^n X_i^{(n)} - \mu_1 \right| \leq \\
& \leq 1(t \leq \theta) \left\{ \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^{[n\theta]} X_i^{(n)} - \frac{([n\theta]-nt)}{n(1-t)} \mu_0 \right| + \right. \\
& + \left. \left| \frac{1}{(n-nt)} \sum_{i=[n\theta]+1}^n X_i^{(n)} - \frac{(n-[n\theta])}{n(1-t)} \mu_1 \right| + \frac{|[n\theta]-n\theta|}{n(1-t)} (\mu_0 + \mu_1) \right\} +
\end{aligned}$$

$$\begin{aligned}
& + 1(t > \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^n X_i^{(n)} - \mu_1 \right| = \\
& = 1(t \leq \theta) \left\{ \frac{([n\theta]-nt)}{n(1-t)} \left| \frac{1}{[n\theta]-nt} \sum_{i=nt+1}^{[n\theta]} X_i^{(n)} - \mu_0 \right| + \right. \\
& \left. + \frac{(n-[n\theta])}{n(1-t)} \left| \frac{1}{n-[n\theta]} \sum_{i=[n\theta]+1}^n X_i^{(n)} - \mu_1 \right| + \frac{|[n\theta]-n\theta|}{n(1-t)} (\mu_0 + \mu_1) \right\} + \\
& + 1(t > \theta) \left| \frac{1}{(n-nt)} \sum_{i=nt+1}^n X_i^{(n)} - \mu_1 \right| = H_2.
\end{aligned}$$

Novamente utilizando (2.3.1.1) e considerando que

$$|[n\theta]-n\theta| \leq 1,$$

$$[n\theta]-nt \leq n-nt,$$

$$n-[n\theta] \leq n-nt, \text{ se } t \leq \theta$$

temos

$$\begin{aligned}
H_2 \leq 1(t \leq \theta) & \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq t \\ l \geq \lceil n\theta \rceil/n}} \left| \frac{1}{n(1-m)} - \sum_{i=nt+1}^{\lceil n\theta \rceil} X_i^{(n)} - \frac{\lceil n\theta \rceil/n - t}{(1-m)} \mu_0 \right| + \right. \\
& + \left. \sup_{\substack{m, l \in \bar{T}_n \\ m \leq \lceil n\theta \rceil}} \left| \frac{1}{n(1-m)} \sum_{i=\lceil n\theta \rceil+1}^n X_i^{(n)} - \frac{(1-\lceil n\theta \rceil/n)}{(1-m)} \mu_1 \right| + \frac{1}{n} \frac{(\mu_0 + \mu_1)}{(1-t)} \right\} + \\
& + 1(t > \theta) \sup_{\substack{m, l \in \bar{T}_n \\ m \leq t \\ l \geq 1}} \left| \frac{1}{n(1-m)} \sum_{i=nt+1}^n X_i^{(n)} - \frac{(1-t)}{(1-m)} \mu_1 \right| = H_t^{(n)*}.
\end{aligned}$$

O que nos fornece uma cota superior para $H_t^{(n)}$.

E, combinando os resultados acima, temos

$$|D_t^{(n)} - \Delta_t| \leq H_t^{(n)} + H_t^{(n)} \leq H_t^{(n)*} + H_t^{(n)*}.$$

e, portanto,

$$|D_t^{(n)*} - \Delta_t^*| \leq t(1-t)(H_t^{(n)*} + H_t^{(n)*}).$$

Como

$$t(1-t) \leq 1/4, \quad t \in T_n,$$

temos que

$$| D_t^{(n)*} - \Delta_t^* | \leq 1/4 ({}_t H^{(n)*} + H_t^{(n)*}). \quad (\text{eq. 2.3.2.1})$$

Por outro lado,

$$\left\{ \sup_{t \in T_n} | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} = \bigcup_{t \in T_n} \left\{ | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\}$$

e, considerando (eq. 2.3.2.1), obtemos

$$\left\{ | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} \subseteq \left\{ 1/4 ({}_t H^{(n)*} + H_t^{(n)*}) > \varepsilon \right\}$$

e

$$\begin{aligned} P \left\{ \sup_{t \in T_n} | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} &= P \left[\bigcup_{t \in T_n} \left\{ | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} \right] \leq \\ &\leq \sum_{t \in T_n} P \left\{ | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} \leq \\ &\leq \sum_{t \in T_n} P \left\{ 1/4 ({}_t H^{(n)*} + H_t^{(n)*}) > \varepsilon \right\} = \end{aligned}$$

$$\begin{aligned}
&= \sum_{t \in T_n} P \left\{ \langle H_t^{(n)*} + H_t^{(n)*} \rangle > 4\varepsilon \right\} = \\
&= \sum_{\substack{t \in T_n \\ t \leq \theta}} P \left\{ \left(\sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq t}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{nt} Y_i^{(n)} - \frac{t}{(1-m)} \mu_0 \right| + \right. \right. \\
&\quad + \sup_{\substack{m, l \in \bar{T}_n \\ m \leq t \\ l \geq [n\theta]/n}} \left| \frac{1}{n(1-m)} - \sum_{i=nt+1}^{[n\theta]} X_i^{(n)} - \frac{([n\theta]/n-t)}{(1-m)} \mu_0 \right| + \\
&\quad + \sup_{\substack{m, l \in \bar{T}_n \\ m \leq [n\theta]/n \\ l \geq 1}} \left| \frac{1}{n(1-m)} \sum_{i=[n\theta]+1}^n X_i^{(n)} - \frac{(1-[n\theta]/n)}{(1-m)} \mu_1 \right| + \\
&\quad \left. + \frac{1}{n} \frac{(\mu_0 + \mu_1)}{(1-t)} \right) > 4\varepsilon \Bigg\} + \\
&+ \sum_{\substack{t \in T_n \\ t > \theta}} P \left\{ \left(\sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq [n\theta]/n}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{[n\theta]} Y_i^{(n)} - \frac{([n\theta]/n)}{(1-m)} \mu_0 \right| + \right. \right. \\
&\quad + \sup_{\substack{m, l \in \bar{T}_n \\ m \leq [n\theta]/n \\ l \geq t}} \left| \frac{1}{n(1-m)} \sum_{i=[n\theta]+1}^{nt} Y_i^{(n)} - \frac{(t-[n\theta]/n)}{(1-m)} \mu_1 \right| +
\end{aligned}$$

$$\begin{aligned}
& + \sup_{\substack{m, l \in \bar{T}_n \\ m \leq l \\ l \geq 1}} \left| \frac{1}{n(1-m)} \sum_{i=nt+1}^n X_i^{(n)} - \frac{(1-t)}{(1-m)} \mu_1 \right| + \\
& + \frac{1}{n} \left\{ \frac{(\mu_0 + \mu_1)}{t} \right\} \Bigg\} > 4\varepsilon \Bigg\} \leq \\
\leq & \sum_{\substack{l \in T_n \\ l \leq \theta}} \left[P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq l}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{nt} Y_i^{(n)} - \frac{t}{(1-m)} \mu_0 \right| > \varepsilon \right\} + \right. \\
& + P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq t \\ l \geq \lceil n\theta \rceil / n}} \left| \frac{1}{n(1-m)} \sum_{i=nt+1}^{\lceil n\theta \rceil} X_i^{(n)} - \frac{(\lceil n\theta \rceil / n - t)}{(1-m)} \mu_0 \right| > \varepsilon \right\} + \\
& + P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq \lceil n\theta \rceil / n \\ l \geq 1}} \left| \frac{1}{n(1-m)} \sum_{i=\lceil n\theta \rceil + 1}^n X_i^{(n)} - \frac{(1 - \lceil n\theta \rceil / n)}{(1-m)} \mu_1 \right| > \varepsilon \right\} + \\
& + P \left\{ \left[\frac{1}{n} \frac{(\mu_0 + \mu_1)}{(1-t)} \right] > \varepsilon \right\} \Bigg] +
\end{aligned}$$

$$\begin{aligned}
& + \sum_{\substack{t \in T_n \\ t > \theta}} \left[P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m=0 \\ l \geq [n\theta]/n}} \left| \frac{1}{n(1-m)} \sum_{i=1}^{[n\theta]} Y_i^{(n)} - \frac{([n\theta]/n)}{(1-m)} \mu_0 \right| > \varepsilon \right\} + \right. \\
& + P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq [n\theta]/n \\ l \geq t}} \left| \frac{1}{n(1-m)} \sum_{i=[n\theta]+1}^{nt} Y_i^{(n)} - \frac{(t-[n\theta]/n)}{(1-m)} \mu_1 \right| > \varepsilon \right\} + \\
& + P \left\{ \sup_{\substack{m, l \in \bar{T}_n \\ m \leq t \\ l \geq 1}} \left| \frac{1}{n(1-m)} \sum_{i=nt+1}^n X_i^{(n)} - \frac{(1-t)}{(1-m)} \mu_1 \right| > \varepsilon \right\} + \\
& + P \left\{ \left(\frac{1}{n} \frac{(\mu_0 + \mu_1)}{t} \right) > \varepsilon \right\} \Bigg].
\end{aligned}$$

Pelo Lema 2.3.1., temos que

$$P \left[\sup_{\substack{m, l \in \bar{T}_n \\ m \leq s \\ l \geq r}} \left| \frac{1}{n(1-m)} \sum_{i=sn+1}^{rn} Y_i^{(n)} - \frac{(r-s)}{(1-m)} \mu \right| > \varepsilon \right] \leq e^{-nJ(\varepsilon)}.$$

Portanto

$$\begin{aligned}
P \left\{ \sup_{t \in T_n} | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} &\leq \sum_{\substack{t \in T_n \\ t \leq \theta}} e^{-nJ^{(1)}(\varepsilon)} + \sum_{\substack{t \in T_n \\ t \leq \theta}} e^{-nJ^{(2)}(\varepsilon)} + \\
&+ \sum_{\substack{t \in T_n \\ t \leq \theta}} e^{-nJ^{(3)}(\varepsilon)} + \sum_{\substack{t \in T_n \\ t \leq \theta}} P \left\{ \left(\frac{1}{n} \frac{(\mu_0 + \mu_1)}{(1-t)} \right) > \varepsilon \right\} + \\
&+ \sum_{\substack{t \in T_n \\ t > \theta}} e^{-nJ^{(4)}(\varepsilon)} + \sum_{\substack{t \in T_n \\ t > \theta}} e^{-nJ^{(5)}(\varepsilon)} + \\
&+ \sum_{\substack{t \in T_n \\ t > \theta}} e^{-nJ^{(6)}(\varepsilon)} + \sum_{\substack{t \in T_n \\ t > \theta}} P \left\{ \left(\frac{1}{n} \frac{(\mu_0 + \mu_1)}{t} \right) > \varepsilon \right\}.
\end{aligned}$$

Como a função $e^{-nJ(\varepsilon)}$ é integrável e as probabilidades são limitadas por 1, temos que cada somatório envolvendo termos exponenciais do limite anterior é finito com cada termo convergindo a zero quando $n \rightarrow \infty$. Nas demais somatórios, cada termo também obviamente é limitado e igual a zero para n suficientemente grande.

Portanto, temos provado que

$$P \left\{ \sup_{t \in T_n} | D_t^{(n)*} - \Delta_t^* | > \varepsilon \right\} \xrightarrow[n \rightarrow \infty]{} 0 \quad \blacksquare$$

Agora, vejamos um outro lema.

Lema 2.3.3. :

$$P \left\{ | \Delta_{\hat{\lambda}_n}^* - \Delta_\theta^* | > \varepsilon \right\} \xrightarrow[n \rightarrow \infty]{} 0.$$

Prova : Seja t_n o maximizador de Δ_t em $T_n = \{ i/n ; i = 1, \dots, n \}$.

Como

$$\Delta_t^* = | \mu_0 - \mu_1 | \left\{ t (1-\theta) 1(t \leq \theta) + (1-t) \theta 1(t > \theta) \right\},$$

t_n é dado por

$$t_n = 1 \left\{ \rho(\lfloor n\theta \rfloor/n) \geq \rho(\lfloor n\theta \rfloor+1/n) \right\} \lfloor n\theta \rfloor/n + \\ + 1 \left\{ \rho(\lfloor n\theta \rfloor/n) < \rho(\lfloor n\theta \rfloor+1/n) \right\} (\lfloor n\theta \rfloor+1)/n.$$

Consideremos as desigualdades

$$\begin{aligned} \left| \Delta_{\hat{\lambda}_n}^* - \Delta_{\theta}^* \right| &\leq \left| \Delta_{\hat{\lambda}_n}^* - D_{\hat{\lambda}_n}^{(n)*} \right| + \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{\theta}^* \right| \\ &\leq \left| \Delta_{\hat{\lambda}_n}^* - D_{\hat{\lambda}_n}^{(n)*} \right| + \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| + \left| \Delta_{t_n}^* - \Delta_{\theta}^* \right|. \end{aligned}$$

Segue que

$$\begin{aligned} P \left[\left| \Delta_{\hat{\lambda}_n}^* - \Delta_{\theta}^* \right| > \varepsilon \right] &\leq P \left[\left| \Delta_{\hat{\lambda}_n}^* - D_{\hat{\lambda}_n}^{(n)*} \right| > \varepsilon/3 \right] + \\ &+ P \left[\left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right] + P \left[\left| \Delta_{t_n}^* - \Delta_{\theta}^* \right| > \varepsilon/3 \right]. \end{aligned}$$

Como

$$\begin{aligned} \Delta_{t_n}^* &= 1 \left\{ \rho([n\theta]/n) \geq \rho([n\theta+1]/n) \right\} \left[|\mu_0 - \mu_1| \rho([n\theta]/n) \right] + \\ &+ 1 \left\{ \rho([n\theta]/n) < \rho([n\theta+1]/n) \right\} \left[|\mu_0 - \mu_1| \rho([n\theta+1]/n) \right] = \\ &= |\mu_0 - \mu_1| \max \left\{ \rho([n\theta]/n), \rho([n\theta+1]/n) \right\}, \end{aligned}$$

$$\Delta_{\theta}^* = |\mu_0 - \mu_1| \theta (1-\theta)$$

e quando $n \longrightarrow \infty$, $[n\theta]/n \longrightarrow \theta$,

temos que

$$\left| \Delta_{t_n}^* - \Delta_{\theta}^* \right| \xrightarrow[n \longrightarrow \infty]{} 0$$

implica em

$$P \left[\left| \Delta_t^* - \Delta_{\theta}^* \right| > \varepsilon/3 \right] \xrightarrow{n \rightarrow \infty} 0.$$

Por outro lado, podemos notar que

$$D_{\hat{\lambda}_n}^{(n)*} \geq \Delta_{t_n}^* \geq \Delta_{\hat{\lambda}_n}^* \quad \text{ou} \quad \Delta_{t_n}^* \geq D_{\hat{\lambda}_n}^{(n)*} \geq D_{t_n}^{(n)*}.$$

Portanto

$$\left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| \leq \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{\hat{\lambda}_n}^* \right| \quad \text{ou} \quad \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| \leq \left| D_{t_n}^{(n)*} - \Delta_{t_n}^* \right|$$

Isto implica que

$$\left\{ \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right\} \subseteq \left\{ \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{\hat{\lambda}_n}^* \right| > \varepsilon/3 \right\} \subseteq \left\{ \sup_{t \in T_n} \left| D_t^{(n)*} - \Delta_t^* \right| > \varepsilon/3 \right\}$$

ou

$$\left\{ \left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right\} \subseteq \left\{ \left| D_{t_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right\} \subseteq \left\{ \sup_{t \in T_n} \left| D_t^{(n)*} - \Delta_t^* \right| > \varepsilon/3 \right\}$$

Portanto, de qualquer forma,

$$P \left[\left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right] \leq P \left[\sup_{t \in T_n} \left| D_t^{(n)*} - \Delta_t^* \right| > \varepsilon/3 \right]$$

e, pelo Lema 2.3.2., temos que

$$P \left[\left| D_{\hat{\lambda}_n}^{(n)*} - \Delta_{t_n}^* \right| > \varepsilon/3 \right] \xrightarrow[n \rightarrow \infty]{} 0.$$

O mesmo lema pode ser aplicado para mostrar que

$$P \left[\left| \Delta_{\hat{\lambda}_n}^* - D_{\hat{\lambda}_n}^{(n)*} \right| > \varepsilon/3 \right] \xrightarrow[n \rightarrow \infty]{} 0.$$

Sendo assim, temos que

$$P \left\{ \left| \Delta_{\hat{\lambda}_n}^* - \Delta_{\theta}^* \right| > \varepsilon \right\} \xrightarrow[n \rightarrow \infty]{} 0. \quad \blacksquare$$

Podemos provar, agora, um importante resultado, que implica na consistência do nosso estimador.

Teorema 2.3.4. : $\hat{\lambda}_n$ é estimador consistente para θ , isto é,

$$P \left[\left| \hat{\lambda}_n - \theta \right| > \varepsilon \right] \xrightarrow[n \rightarrow \infty]{} 0.$$

Prova :

$$\left| t - \theta \right| > \nu$$

implica em

$$\left| \Delta_t^* - \Delta_{\theta}^* \right| = \left| \mu_0 - \mu_1 \right| \left\{ t(1-\theta) 1_{\{t \leq \theta\}} + (1-t)\theta 1_{\{t > \theta\}} - \theta(1-\theta) \right\} =$$

$$\begin{aligned}
&= | \mu_0 - \mu_1 | \left\{ (1-\theta) (1-\theta) 1\{t \leq \theta\} + (1-t-1+\theta) \theta 1\{t > \theta\} \right\} > \\
&> | \mu_0 - \mu_1 | \left\{ \nu (1-\theta) 1\{t \leq \theta\} + \nu \theta 1\{t > \theta\} \right\} = \\
&= \nu | \mu_0 - \mu_1 | \left\{ (1-\theta) 1\{t \leq \theta\} + \theta 1\{t > \theta\} \right\} = \nu K \quad (K > 0).
\end{aligned}$$

Portanto

$$P \left[| \hat{\lambda}_n - \theta | > \varepsilon \right] \leq P \left[| \Delta_{\hat{\lambda}_n}^* - \Delta_{\theta}^* | > K\varepsilon \right]$$

e, pelo Lema 2.3.3., temos que

$$P \left\{ | \Delta_{\hat{\lambda}_n}^* - \Delta_{\theta}^* | > K\varepsilon \right\} \xrightarrow[n \rightarrow \infty]{} 0$$

Então

$$P \left[| \hat{\lambda}_n - \theta | > \varepsilon \right] \xrightarrow[n \rightarrow \infty]{} 0. \quad \blacksquare$$

Portanto, temos provada o Teorema 2.3.1., isto é, temos que, neste caso de independência, nosso estimador é consistente.

2.4. Aplicação do estimador nos dados do Rio Nilo.

O exemplo de aplicação da vazão do Rio Nilo citado na introdução do capítulo anterior é clássico em termos de trabalhos abordando o problema de ponto de mudança, considerando sequência de variáveis aleatórias independentes.

A tabela abaixo apresenta os dados contidos em Cobb (1978). É importante aqui lembrarmos que o trabalho cita o fato de haver evidências meteorológicas que a mudança no volume anual de descarga do rio ocorreu em 1898.

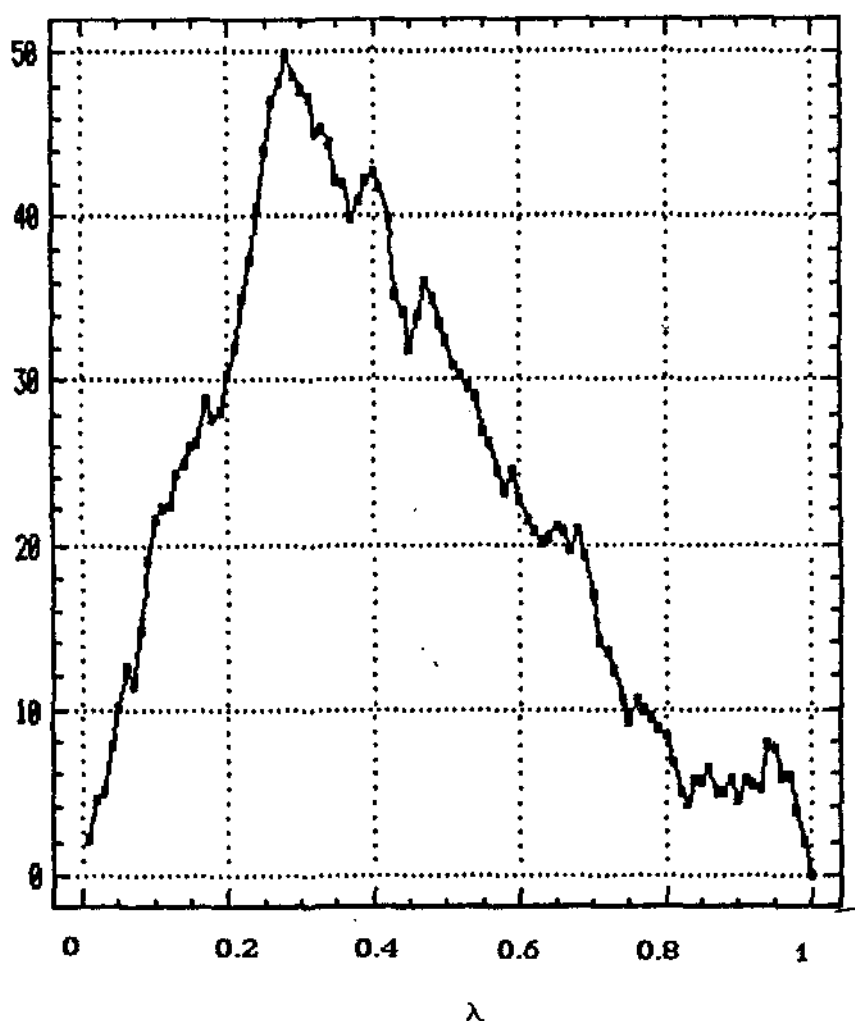
Tabela 2.4.1. Volume anual de descarga (desc) do Rio Nilo (10^9 m^3).

De 1871 a 1970,

ano	desc	ano	desc	ano	desc	ano	desc
1871	1120	1896	1220	1921	768	1946	1040
1872	1160	1897	1030	1922	845	1947	860
1873	963	1898	1100	1923	864	1948	874
1874	1210	1899	774	1924	862	1949	848
1875	1160	1900	840	1925	698	1950	890
1876	1160	1901	874	1926	845	1951	744
1877	813	1902	694	1927	744	1952	749
1878	1230	1903	940	1928	796	1953	838
1879	1370	1904	833	1929	1040	1954	1050
1880	1140	1905	701	1930	759	1955	918
1881	995	1906	916	1931	781	1956	986
1882	935	1907	692	1932	865	1957	797
1883	1110	1908	1020	1933	845	1958	923
1884	994	1909	1050	1934	944	1959	975
1885	1020	1910	969	1935	984	1960	815
1886	960	1911	831	1936	897	1961	1020
1887	1180	1912	726	1937	822	1962	906
1888	799	1913	456	1938	1010	1963	901
1889	958	1914	824	1939	771	1964	1170
1890	1140	1915	702	1940	676	1965	912
1891	1100	1916	1120	1941	649	1966	746
1892	1210	1917	1100	1942	846	1967	919
1893	1150	1918	832	1943	812	1968	718
1894	1250	1919	764	1944	742	1969	714
1895	1260	1920	821	1945	801	1970	740

Abaixo, temos o gráfico mostrando a série correspondente à função $D_{\lambda}^{(n)*}$, onde observamos como ponto de máximo o valor de λ igual a 0.28, o que nos indica o ano de 1898 para o ponto de mudança da média do volume de descarga do Rio Nilo.

Gráfico 2.4.1 Função amostral $D_{\lambda}^{(n)*}$.



2.5. Proposta de um segundo estimador

Nos estudos de simulação Monte Carlo que será apresentado na próxima seção, foi incorporado um novo estimador que, dada a sua complexidade, não nos foi possível chegar a alguma conclusão sobre suas propriedades teoricamente. A motivação para este segundo estimador foi baseada no fato do primeiro ($\hat{\lambda}_n$) apresentar uma assimetria bastante acentuada na variabilidade em relação ao valor do parâmetro, fato que observamos nos estudos de simulação. Ver gráficos no apêndice B.

Então foi feita uma correção estocástica sobre $\hat{\lambda}_n$ que define o novo estimador, denotado por $\hat{\lambda}_n^{*(e)}$, como o valor λ tal que

$$D_{\hat{\lambda}_n^{*(e)}}^n = \max_{\lambda \in T_n} D_{\lambda}^n$$

onde

$$D_{\lambda}^n = \frac{\lambda(1-\lambda) |\bar{X}^n - \bar{X}_{\lambda}^n|}{\sqrt{\frac{\lambda S^2}{n\lambda} + \frac{S_{\lambda}^2}{(n-n\lambda)}}},$$

$$\lambda S^2 = \frac{\sum_{i=1}^{n\lambda} (X_i^n - \bar{X}^n)^2}{n\lambda - 1}$$

e

$$S_{\lambda}^2 = \frac{\sum_{i=n\lambda+1}^n (X_i^n - \bar{X}_{\lambda}^n)^2}{n - n\lambda - 1}.$$

2.6. Estudos de simulação.

Para este estudo que segue, foi utilizado, para a geração de números aleatórios, o gerador congruencial linear, cujo algoritmo está apresentado no Apêndice C deste trabalho (Referência: Bustos (1992)).

Foram feitas 2000 simulações, onde foram estudados os comportamentos dos seguintes estimadores: o de máxima verossimilhança proposto por Hinkley (1970) $(\hat{\theta}_1)$, $\hat{\lambda}_n(\hat{\theta}_2)$ e $\hat{\lambda}_n^{(*)}(\hat{\theta}_3)$ propostos e estudados no presente trabalho.

Foram calculados as medianas, os percentis de 5% e 95 %, os erros quadráticos médios e os respectivos intervalos de confiança. A cada caso simulado foi fornecida a semente de geração de números aleatórios (s0).

Foram estudados o caso Normal com variância constante e igual a um e o caso Bernoulli para diferentes mudanças na média, diferentes valores de θ e diferentes tamanhos de amostras (n). Estes casos foram citados e estudados nos trabalhos de Hinkley (1970) e Hinkley & Hinkley (1970).

As tabelas com os resultados se encontram no Apêndice A (Tabelas 1, 2, 3 e 4), os gráficos das distribuições de frequências, no caso bernoulli, no Apêndice B e os programas no Apêndice D (Programas 1 e 2).

No caso normal, pode-se notar que, quando n cresce, as medianas dos três estimadores indicam que estão convergindo para o

valor de θ , sendo que $\hat{\theta}_1$ com uma maior velocidade quando $\theta = 0.2$. Por outro lado, podemos perceber, observando os percentis de 5% e 95% que a amplitude entre eles é notadamente maior para $\hat{\theta}_1$, o que também se confirma analisando os erros médios quadráticos, isto é, há uma maior variabilidade no estimador de máxima verossimilhança do que no estimador proposto neste trabalho. Isto se torna mais visível quanto maior o tamanho da amostra e maior a diferença entre as médias pré e pós. No caso $\theta = 0.5$, acontecem as mesmas coisas, só que de uma maneira mais notável.

No caso Bernoulli, podemos observar o mesmo comportamento (Tabelas 3 e 4) dos três estimadores que no caso Normal e, pelos gráficos (Apêndice B), podemos comparar os comportamentos dos dois estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$, onde observamos que, na maioria dos casos, são bastante similares.

Além destas, consideramos outras distribuições discretas em nossos estudos de simulação. Foram utilizadas duas combinações de distribuições (pré (F) e pós (G) ponto de mudança), as mesmas utilizadas no trabalho de Perez-Abreu (1989) para comparar o estimador proposto no mesmo com o proposto por Carlstein (1988). São elas :

Caso (1) : F ~ Binomial(10,0.1) e G ~ Binomial(10,0.3).

(Tabela 5 - Apêndice A).

Caso (2) : F ~ Poisson(2) e G ~ Geométrica (0.5).

(Tabela 6 - Apêndice B).

Foram feitas 1000 simulações para tamanhos de amostras (n)

iguais a 25, 50 e 100, calculados a média do estimador $\hat{\lambda}_n$ ($\hat{\theta}_1$), o erro médio absoluto e os respectivos intervalos de confiança. Além destes resultados, foram reapresentados os resultados de Perez-Abreu (1989) para o estimador proposto pelo mesmo ($\hat{\theta}_2$) e o de Carlstein (1988) ($\hat{\theta}_3$).

Pela Tabela 5, podemos notar que, no Caso (1), nosso estimador tem um menor erro médio absoluto, apesar de não termos informação sobre o intervalo de confiança para este erro nos dois últimos estimadores, quando θ está próximo de 0.5. Agora, apesar de ser pior quando θ está próximo de 0.9, é bastante razoável para ser considerado, dado a sua simplicidade.

Na Tabela 6, que mostra os resultados do Caso (2), observamos comportamento similar.

2.7. Conclusões.

Nos estudos de simulação, pudemos observar, através dos gráficos apresentados no Apêndice B, que o estimador proposto ($\hat{\lambda}_n$) tem uma convergência para o verdadeiro valor de θ que talvez, se aumentássemos mais o tamanho da amostra (maior que 100), ficaria cada vez mais notável e se destacaria em relação ao estimador de máxima verossimilhança, o que é apenas um reflexo de sua propriedade de ser consistente.

Quanto à comparação com os estimadores propostos por Carlstein

(1988) e Perez-Abreu (1989), temos que o estimador proposto neste trabalho tem um comportamento de igual a melhor quando existe a mudança na média. Devemos levar em consideração, entretanto, que os outros dois estimadores tem uma maior abrangência, principalmente o de Carlstein, que pode ser aplicado tanto a distribuições discretas quanto a contínuas e considerando qualquer mudança na distribuição.

O comportamento do estimador $\hat{\lambda}_n$ proposto neste trabalho mostra certas propriedades que demonstram ser suficientemente bom para ser considerado. Devemos levar em conta sua simplicidade e abrangência sobre diversos aspectos como : fácil manejo, aplicado tanto a distribuições contínuas quanto discretas, não necessitar conhecimento prévio da distribuição e sua consistência.

Devemos ressaltar que a única suposição que deve ser satisfeita para que o estimador seja consistente é que ele tenha função geradora de momentos $M(t)$ finita para todo t . Cabe ressaltar uma vez mais que esta suposição é satisfeita pelas distribuições usuais, como normal, binomial, exponencial, poisson, geométrica, etc.

CAPÍTULO 3.

APLICAÇÃO DO ESTIMADOR EM PROCESSOS DE RAMIFICAÇÃO.

Neste capítulo, vamos aplicar o estimador proposto no capítulo anterior para estimar o ponto de mudança da média do número de descendentes em um processo de ramificação.

3.1. Introdução.

Nesta seção, vamos introduzir conceitos básicos em Processos de Ramificação, um tipo particular de Cadeias de Markov. Iremos nos restringir ao caso de tempo discreto.

Definição 3.1.1. Suponhamos ξ_1, ξ_2, \dots uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) assumindo valores em $S = \{0, 1, 2, \dots\}$ e tendo distribuição F . Uma cadeia de Markov é chamada um processo de ramificação com espaço de estados S se suas probabilidades de transição são dadas por

$$f(x, y) = P\{X_t = y \mid X_{t-1} = x\} = P\{\xi_1 + \dots + \xi_x = y\}$$

para $x \geq 1$, $y \geq 0$ e $p(0,0) = 1$.

Ou seja,

$$\begin{aligned} X_0 &= x_0, \\ X_1 &= \xi_1^{(0)} + \dots + \xi_{x_0}^{(0)}, \\ X_2 &= \xi_1^{(1)} + \dots + \xi_{X_1}^{(1)}, \\ &\vdots \\ X_t &= \xi_1^{(t-1)} + \dots + \xi_{X_{t-1}}^{(t-1)}, \\ &\vdots \end{aligned}$$

onde $\xi_1^{(0)}, \xi_2^{(0)}, \dots, \xi_1^{(1)}, \xi_2^{(1)}, \dots, \xi_1^{(t)}, \xi_2^{(t)}, \dots$ são variáveis aleatórias i.i.d. com distribuição F. $\xi_i^{(t)}$ pode ser interpretado como o número de descendentes do i -ésimo indivíduo da t -ésima geração e X_t é o número total de indivíduos da t -ésima geração.

Definição 3.1.2. Probabilidades de extinção.

Um aspecto de grande importância em processos de ramificação é a probabilidade da população se extinguir, que é dada por

$$P(X_m = 0 \text{ para algum } m) = P\left(\lim_{m \rightarrow \infty} [X_m = 0]\right).$$

Com respeito a esta probabilidade, um resultado bastante importante é dado no teorema seguinte.

Teorema 3.1.1. Sejam $X = \langle X_t ; t \in \mathbb{N} \rangle$ um processo de ramificação, μ a média do número de descendentes de um certo indivíduo, isto é, $\mu = E(\xi)$ e ρ a probabilidade de extinção do processo quando $X_0 = 1$. Temos então que

$$(i) \quad \rho = 1 \iff \mu \leq 1,$$

$$(ii) \quad \rho < 1 \iff \mu > 1.$$

Prova : Ver Karlin and Taylor (1985).

O teorema diz que a extinção é certa se e somente se o número médio de descendentes por indivíduo não excede um.

E é por este motivo que é de grande relevância, num processo de ramificação, conhecermos o valor de μ e sermos capazes de estimar tempos de mudança deste parâmetro, caso estes ocorram.

3.2. Estimação não-paramétrica do ponto de mudança de distribuição em processos de ramificação.

Como pudemos observar na seção anterior, a média do número de descendentes de um indivíduo (μ) caracteriza o processo de ramificação com relação a probabilidade de extinção. Considerando este fato, o problema de mudança de distribuição, neste contexto, poderia ficar restrito ao de mudança no primeiro momento da distribuição.

Consideremos um processo de ramificação $X = (X_t : t \in \mathbb{N})$, onde o número de descendentes do r -ésimo indivíduo da t -ésima geração até um tempo t_0 , é uma variável aleatória $\xi_r^{(t)}$ com uma distribuição F e após o tempo t_0 , é uma variável aleatória que denotaremos por $\eta_r^{(t)}$, também assumindo valores em $\mathbb{Z} = \{0, 1, \dots\}$, com uma outra distribuição G . Então, temos que X satisfaz as condições descritas a seguir :

$$X_0 = 1,$$

$$X_{t+1} = \begin{cases} \sum_{r=1}^{X_t} \xi_r^{(t)} & , \quad t \leq t_0 \\ \sum_{r=1}^{X_t} \eta_r^{(t)} & , \quad t > t_0 \end{cases}$$

com

$$E(\xi_r^{(t)}) = \mu_\xi, \quad E(\eta_r^{(t)}) = \mu_\eta,$$

onde

$$\mu_\xi \neq \mu_\eta.$$

Agora, vamos considerar $X^{(n)} = (X_1^{(n)}, X_2^{(n)}, \dots, X_n^{(n)})$, uma observação do processo de ramificação X . Vamos assumir que $t_0 \leq n$ e que não ocorra extinção do processo até a $(n-1)$ -ésima geração, isto é, $X_{n-1} > 0$. Sendo assim, passamos a considerar o problema fazendo $t_0 = \lfloor n\theta \rfloor$, com $\theta \in [0,1]$. Portanto, nosso interesse, como no capítulo anterior, passa a ser

estimar θ .

O fato de supormos a não extinção do processo até a n -ésima geração é bastante natural, dado que, ao observarmos um processo, sempre iremos considerar um determinado número de gerações onde não ocorrerá a extinção.

Para que possamos estimar o ponto de mudança da média do número de descendentes, precisamos de um estimador para esta média. Mas acontece que não observamos o número de descendentes por indivíduo, mas sim o número total de indivíduos a cada geração.

A seguir, temos uma proposição que nos auxilia na solução deste problema.

Proposição 3.2.1. Seja X um processo de ramificação com média do número de descendentes μ . Então, dado $X_t > 0$,

$$\frac{X_{t+1}}{X_t}$$

é estimador não viciado para μ .

Prova : Temos que

$$E \left(X_{t+1} / X_t \mid X_t > 0 \right) = E \left(E \left(X_{t+1} / X_t \mid X_t > 0, X_t \right) \right).$$

Mas

$$E \left(X_{t+1} / X_t \mid X_t > 0, X_t = k \right) =$$

$$\begin{aligned}
&= E \left(\sum_{r=1}^k \xi_r^{(t)} / k \mid X_t > 0, X_t = k \right) = \\
&= E \left(\sum_{r=1}^k \xi_r^{(t)} / k \right) = \sum_{r=1}^k E \left(\xi_r^{(t)} \right) / k = \mu.
\end{aligned}$$

Portanto,

$$E \left(X_{t+1} / X_t \mid X_t > 0 \right) = E \left(\mu \right) = \mu.$$

E segue que o estimador é não viciado. ■

O estimador acima considerado é denominado **estimador de razão** em Nanthi (1984).

A partir disso, assumindo a não extinção do processo de ramificação até a n -ésima geração, vamos considerar uma nova sequência de variáveis aleatórias

$$Y^{(n)} = \{ Y_t^{(n)} ; t \in \{ 1, \dots, n \} \},$$

onde

$$Y_t^{(n)} = \frac{X_{t+1}^{(n)}}{X_t^{(n)}},$$

com $X_t^{(n)} > 0$.

Vamos assumir que o segundo momento da variável $Y_t^{(n)}$ é finito daqui para frente.

Temos, então, a seguinte proposição :

proposição 3.2.2. Consideremos $Y^{(n)}$ uma sequência de variáveis aleatórias, onde

$$Y_t^{(n)} = \frac{X_{t+1}^{(n)}}{X_t^{(n)}},$$

com $X_t^{(n)} > 0$, associadas ao processo de ramificação X . Suponhamos que ocorra uma mudança na média do número de descendentes no ponto $\ln \theta$. Então temos que

$$i) E \langle Y_t^{(n)} \mid X_t^{(n)} > 0 \rangle = \begin{cases} \mu_\xi & , t \leq \ln \theta \\ \mu_\eta & , t > \ln \theta \end{cases};$$

$$ii) \text{Cov} \langle Y_t^{(n)}, Y_s^{(n)} \mid X_t^{(n)} > 0, X_s^{(n)} > 0 \rangle = 0, \quad t \neq s;$$

$$iii) \text{Var} \langle Y_t^{(n)} \mid X_t^{(n)} > 0 \rangle = \begin{cases} \sigma_\xi^2 E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right], & t \leq \ln \theta \\ \sigma_\eta^2 E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right], & t > \ln \theta \end{cases}$$

prova :

$$i) E \langle Y_t^{(n)} \mid X_t^{(n)} > 0 \rangle = E \langle E \langle Y_t^{(n)} \mid X_t^{(n)} > 0, X_t^{(n)} \rangle \rangle.$$

as

$$\langle Y_t^{(n)} \mid X_t^{(n)} > 0, X_t^{(n)} = x_t^{(n)} \rangle =$$

$$= \begin{cases} E \left[\frac{\sum_{i=1}^{x_t^{(n)}} \xi_i^{(t)}}{x_t^{(n)}} \mid X_t^{(n)} > 0, X_t^{(n)} = x_t^{(n)} \right] = \frac{\sum_{i=1}^{x_t^{(n)}} E(\xi_i^{(t)})}{x_t^{(n)}} = \mu_\xi, & t \leq [n\theta] \\ E \left[\frac{\sum_{i=1}^{x_t^{(n)}} \eta_i^{(t)}}{x_t^{(n)}} \mid X_t^{(n)} > 0, X_t^{(n)} = x_t^{(n)} \right] = \frac{\sum_{i=1}^{x_t^{(n)}} E(\eta_i^{(t)})}{x_t^{(n)}} = \mu_\eta, & t > [n\theta]. \end{cases}$$

Portanto

$$E \langle Y_t^{(n)} \mid X_t^{(n)} > 0 \rangle = \begin{cases} \mu_\xi & , \quad t \leq [n\theta], \\ \mu_\eta & , \quad t > [n\theta]. \end{cases}$$

(ii) Temos também que

$$\begin{aligned} E \langle Y_{t_1}^{(n)} Y_{t_2}^{(n)} \mid X_{t_1}^{(n)} > 0, X_{t_2}^{(n)} > 0 \rangle &= \\ &= E \langle E \langle Y_{t_1}^{(n)} Y_{t_2}^{(n)} \mid X_{t_1}^{(n)} > 0, X_{t_2}^{(n)} > 0, X_{t_1}^{(n)} = x_{t_1}^{(n)}, X_{t_2}^{(n)} = x_{t_2}^{(n)} \rangle \rangle. \end{aligned}$$

Então, para $t_1 \leq [n\theta]$ e $t_2 \leq [n\theta]$

$$\begin{aligned} E \langle Y_{t_1}^{(n)} Y_{t_2}^{(n)} \mid X_{t_1}^{(n)} > 0, X_{t_2}^{(n)} > 0, X_{t_1}^{(n)} = x_{t_1}^{(n)}, X_{t_2}^{(n)} = x_{t_2}^{(n)} \rangle &= \\ &= E \left[\frac{\sum_{i=1}^{x_{t_1}^{(n)}} \xi_i^{(t_1)}}{x_{t_1}^{(n)}} \frac{\sum_{i=1}^{x_{t_2}^{(n)}} \xi_i^{(t_2)}}{x_{t_2}^{(n)}} \mid X_{t_1}^{(n)} > 0, X_{t_2}^{(n)} > 0 \right] = \end{aligned}$$

$$E \left(\frac{\sum_{i=1}^{x_{t_1}^{(n)}} \xi_i^{(t_1)}}{x_{t_1}^{(n)}} \mid X_{t_1}^{(n)} > 0 \right) E \left(\frac{\sum_{i=1}^{x_{t_2}^{(n)}} \xi_i^{(t_2)}}{x_{t_2}^{(n)}} \mid X_{t_2}^{(n)} > 0 \right) = \mu_{\xi}^2.$$

Segue que

$$E \left(Y_{t_1}^{(n)} Y_{t_2}^{(n)} \mid X_{t_1}^{(n)} > 0 \right) = \mu_{\xi}^2$$

Analogamente, encontramos o resultado para os outros casos e obtemos

$$E \left(Y_{t_1}^{(n)} Y_{t_2}^{(n)} \mid X_{t_1}^{(n)} > 0, X_{t_2}^{(n)} > 0 \right) = \begin{cases} \mu_{\xi}^2 & , \quad t_1, t_2 \leq [n\theta], \\ \mu_{\eta}^2 & , \quad t_1, t_2 > [n\theta], \\ \mu_{\xi} \mu_{\eta} & , \quad t_1 \leq [n\theta], t_2 > [n\theta]. \end{cases}$$

Portanto, temos que $Y_{t_1}^{(n)}$ e $Y_{t_2}^{(n)}$ são não correlacionadas.

(iii) Vejamos, agora, o que acontece com a variância das variáveis aleatórias $Y_t^{(n)}$'s deste processo :

Caso 1 : $t \leq [n\theta]$

$$\text{Var} \left(Y_t^{(n)} \mid X_t^{(n)} > 0 \right) = E \left(Y_t^{(n)2} \mid X_t^{(n)} > 0 \right) - E^2 \left(Y_t^{(n)} \mid X_t^{(n)} > 0 \right).$$

Mas

$$E \langle Y_t^{(n)^2} \mid X_t^{(n)} > 0 \rangle = E \langle E \langle Y_t^{(n)^2} \mid X_t^{(n)} > 0, X_t^{(n)} \rangle \rangle.$$

Agora

$$\begin{aligned} E \langle Y_t^{(n)^2} \mid X_t^{(n)} > 0, X_t^{(n)} = x_t^{(n)} \rangle &= E \left[\left\langle \left[\frac{\sum_{i=1}^{x_t^{(n)}} \xi_i^{(t)}}{x_t^{(n)}} \right]^2 \mid X_t^{(n)} > 0 \right\rangle \right] = \\ &= \frac{1}{x_t^{(n)^2} } \left[\text{Var} \left[\sum_{i=1}^{x_t^{(n)}} \xi_i^{(t)} \right] + E^2 \left[\sum_{i=1}^{x_t^{(n)}} \xi_i^{(t)} \right] \right] = \\ &= \frac{1}{x_t^{(n)^2} } \left[x_t^{(n)} \sigma_\xi^2 + x_t^{(n)^2} \mu_\xi^2 \right] = \frac{1}{x_t^{(n)}} \sigma_\xi^2 + \mu_\xi^2. \end{aligned}$$

Segue que

$$E \langle Y_t^{(n)^2} \rangle = \mu_\xi^2 + \sigma_\xi^2 E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right]$$

e, portanto

$$\boxed{\text{Var} \langle Y_t^{(n)} \mid X_t^{(n)} > 0 \rangle = \sigma_\xi^2 E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right].}$$

Caso 2 : $t > \ln \theta$

Analogamente ao Caso 1, obtemos

$$\text{Var} \left(Y_t^{(n)} \mid X_t^{(n)} > 0 \right) = \sigma_\eta^2 \ E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right].$$

Combinando os casos 1 e 2, temos que, dado $X_t^{(n)} > 0$,

$$\text{Var}(Y_t^{(n)}) = \sigma_\xi^2 \ E \left[\frac{1}{X_t^{(n)}} \right] 1\{t \leq [n\theta]\} + \sigma_\eta^2 \ E \left[\frac{1}{X_t^{(n)}} \right] 1\{t > [n\theta]\}.$$

Vamos aplicar, agora, o estimador $\hat{\lambda}_n$ do capítulo anterior nesta sequência de variáveis aleatórias não correlacionadas. Notemos também que não ocorre a independência como no capítulo anterior.

Temos que $\hat{\lambda}_n$ é o valor de λ que maximiza a função

$$D_\lambda^{(n)*} = \lambda (1-\lambda) \mid \lambda \bar{Y}^{(n)} - \bar{Y}_\lambda^{(n)} \mid$$

3.3. Estudo das propriedades do estimador.

Ocorre, neste caso, que a variância de $d_\lambda^{(n)*}$ será dada em termos de $E \left[\frac{1}{X_t^{(n)}} \mid X_t^{(n)} > 0 \right]$, o que torna complexo a verificação das propriedades do estimador. Vamos calcular, então, a variância da diferença $\lambda \bar{Y}^{(n)} - \bar{Y}_\lambda^{(n)}$.

Se $Y_t^{(n)}$, s são não correlacionadas, $\lambda \bar{Y}^{(n)}$ e $\bar{Y}_\lambda^{(n)}$ também são não correlacionadas. Então, dado que $X_t^{(n)} > 0, \dots, X_{n-1}^{(n)} > 0$,

$$\text{Var} \left(\frac{1}{\lambda} \bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)} \right) = \text{Var} \left(\frac{1}{\lambda} \bar{Y}^{(n)} \right) + \text{Var} \left(\bar{Y}_{\lambda}^{(n)} \right).$$

Mas

Caso 1: $\lambda \leq \theta$

$$\begin{aligned} \text{Var} \left(\frac{1}{\lambda} \bar{Y}^{(n)} \right) &= \text{Var} \left(\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} Y_i^{(n)} \right) = \frac{1}{[n\lambda]^2} \left[\sum_{i=1}^{[n\lambda]} \text{Var} \left(Y_i^{(n)} \right) \right] = \\ &= \frac{1}{[n\lambda]^2} \left[\sum_{i=1}^{[n\lambda]} \left(\sigma_{\xi}^2 E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right) \right] = \\ &= \frac{\sigma_{\xi}^2}{[n\lambda]^2} \sum_{i=1}^{[n\lambda]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right). \end{aligned}$$

$$\begin{aligned} \text{Var}(\bar{Y}_{\lambda}^{(n)}) &= \text{Var} \left(\frac{1}{(n-[n\lambda])} \sum_{i=[n\lambda]+1}^n Y_i^{(n)} \right) = \frac{1}{(n-[n\lambda])^2} \left[\sum_{i=[n\lambda]+1}^n \text{Var} \left(Y_i^{(n)} \right) \right] = \\ &= \frac{1}{(n-[n\lambda])^2} \left[\text{Var} \left(\sum_{i=[n\lambda]+1}^{[n\theta]} Y_i^{(n)} \right) + \text{Var} \left(\sum_{i=[n\theta]+1}^n Y_i^{(n)} \right) \right] = \\ &= \frac{1}{(n-[n\lambda])^2} \left[\sum_{i=[n\lambda]+1}^{[n\theta]} \left(\sigma_{\xi}^2 E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right) + \right. \\ &\quad \left. + \sum_{i=[n\theta]+1}^n \left(\sigma_{\eta}^2 E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right) \right] = \\ &= \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\xi}^2 \sum_{i=[n\lambda]+1}^{[n\theta]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) + \right. \end{aligned}$$

$$+ \sigma_{\eta}^2 \sum_{i=[n\theta]+1}^n E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \Bigg].$$

Portanto

$$\begin{aligned} \text{Var}_{\lambda}(\bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)}) &= \frac{\sigma_{\xi}^2}{[n\lambda]^2} \sum_{i=1}^{[n\lambda]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] + \\ &+ \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\xi}^2 \sum_{i=[n\lambda]+1}^{[n\theta]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right. \\ &\left. + \sigma_{\eta}^2 \sum_{i=[n\theta]+1}^n E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right]. \end{aligned}$$

Caso 2 : $\lambda > \theta$

$$\begin{aligned} \text{Var}_{\lambda}(\bar{Y}^{(n)}) &= \text{Var} \left[\frac{1}{[n\lambda]} \sum_{i=1}^{[n\lambda]} Y_i^{(n)} \right] = \text{Var} \left[\frac{1}{[n\lambda]} \left(\sum_{i=1}^{[n\theta]} Y_i^{(n)} + \sum_{i=[n\theta]+1}^{[n\lambda]} Y_i^{(n)} \right) \right] = \\ &= \frac{1}{[n\lambda]^2} \left[\sum_{i=1}^{[n\theta]} \text{Var} \left(Y_i^{(n)} \right) + \sum_{i=[n\theta]+1}^{[n\lambda]} \text{Var} \left(Y_i^{(n)} \right) \right] = \\ &= \frac{1}{[n\lambda]^2} \left[\sum_{i=1}^{[n\theta]} \left(\sigma_{\xi}^2 E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right) + \right. \\ &\quad \left. + \sum_{i=[n\theta]+1}^{[n\lambda]} \left(\sigma_{\eta}^2 E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right) \right] = \end{aligned}$$

$$= \frac{1}{[n\lambda]^2} \left[\sigma_{\xi}^2 \sum_{i=1}^{[n\theta]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) + \right. \\ \left. + \sigma_{\eta}^2 \sum_{i=[n\theta]+1}^{[n\lambda]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right].$$

$$\text{Var}(\bar{Y}_{\lambda}^{(n)}) = \text{Var} \left(\frac{1}{n-[n\lambda]} \sum_{i=[n\lambda]+1}^n Y_i^{(n)} \right) = \frac{1}{(n-[n\lambda])^2} \left[\sum_{i=[n\lambda]+1}^n \text{Var} (Y_i^{(n)}) \right] = \\ = \frac{1}{(n-[n\lambda])^2} \left[\sum_{i=[n\lambda]+1}^n \left(\sigma_{\eta}^2 E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right) \right] = \\ = \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\eta}^2 \sum_{i=[n\lambda]+1}^n E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right].$$

Portanto

$$\text{Var}(\zeta_{\lambda} \bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)}) = \frac{1}{[n\lambda]^2} \left[\sigma_{\xi}^2 \sum_{i=1}^{[n\theta]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) + \right. \\ \left. + \sigma_{\eta}^2 \sum_{i=[n\theta]+1}^{[n\lambda]} E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right] + \\ + \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\eta}^2 \sum_{i=[n\lambda]+1}^n E \left(\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right) \right].$$

Unindo caso 1 com caso 2, temos que

$$\text{Var} (d_{\lambda}^{(n)}) = \text{Var} (\zeta_{\lambda} \bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)}) =$$

$$1_{Q \leq \theta} \left\{ \frac{\sigma_\xi^2}{[\ln \lambda]^2} \sum_{i=1}^{[n\lambda]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right\}$$

$$\frac{1}{(n - [\ln \lambda])^2} \left[\sigma_\xi^2 \sum_{i=[n\lambda]+1}^{[n\theta]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right]$$

$$\left. \frac{\sigma_\eta^2}{\eta} \sum_{i=[n\theta]+1}^n E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right] \right\}$$

$$1_{Q > \theta} \left\{ \frac{1}{[\ln \lambda]^2} \left[\sigma_\xi^2 \sum_{i=1}^{[n\theta]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] + \sigma_\eta^2 \sum_{i=[n\theta]+1}^{[n\lambda]} E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right] \right\}$$

$$\frac{1}{(n - [\ln \lambda])^2} \left[\sigma_\eta^2 \sum_{i=[n\lambda]+1}^n E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \right] \right\}.$$

Para $d_\lambda^{(n)}$ e f_λ como no capítulo anterior, isto é, $d_\lambda^{(n)} = \bar{Y}_\lambda - \bar{Y}_\lambda$

$$\begin{aligned} E(d_\lambda^{(n)}) &= \left\{ \mu_0 - \frac{1}{n - [\ln \lambda]} \left[([\ln \theta] - [\ln \lambda]) \mu_0 + (n - [\ln \theta]) \mu_1 \right] \right\} 1_{Q \leq \theta} \\ &\quad + \left\{ \frac{1}{[\ln \lambda]} \left[[\ln \theta] \mu_0 + ([\ln \lambda] - [\ln \theta]) \mu_1 \right] - \mu_1 \right\} 1_{Q > \theta}, \end{aligned}$$

temos a seguinte proposição :

proposição 3.2.3. $d_\lambda^{(n)}$ é estimador consistente para f_λ .

prova : Sob a suposição que $X_i > 1$, temos

$$E \left[\frac{1}{X_i^{(n)}} \mid X_i^{(n)} > 0 \right] \leq 1 \quad (\forall i).$$

segue que

$$\text{Var} (d_{\lambda}^{(n)}) \leq$$

$$\begin{aligned} &\leq 1_{0 \leq \theta} \left\{ \frac{\sigma_{\xi}^2}{[n\lambda]^2} [n\lambda] + \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\xi}^2 ([n\theta] - [n\lambda]) + \right. \right. \\ &\quad \left. \left. + \sigma_{\eta}^2 (n - [n\theta]) \right] \right\} + 1_{0 > \theta} \left\{ \frac{1}{[n\lambda]^2} \left[\sigma_{\xi}^2 [n\theta] + \right. \right. \\ &\quad \left. \left. + \sigma_{\eta}^2 ([n\lambda] - [n\theta]) \right] + \frac{1}{(n-[n\lambda])^2} \left[\sigma_{\eta}^2 (n - [n\lambda]) \right] \right\}. \end{aligned}$$

Como podemos observar, o termo n^2 está sempre no denominador, portanto, esta variância tende a zero quando n tende a infinito. Temos, então, utilizando a desigualdade de Tchebychev, que

$$P (| d_{\lambda}^{(n)} - f_{\lambda} | > \varepsilon) \leq \text{Var} (d_{\lambda}^{(n)}) / \varepsilon^2 \xrightarrow[n \rightarrow \infty]{} 0.$$

Portanto, temos que $d_{\lambda}^{(n)}$ é consistente para f_{λ} . ■

Agora, recordando, temos que

$$D_{\lambda}^{(n)*} = | d_{\lambda}^{(n)*} | = | \lambda (1-\lambda) d_{\lambda}^{(n)} |$$

e, portanto,

$$\begin{aligned} \text{Var} \left[d_{\lambda}^{(n)*} \right] &= \text{Var} \left[\lambda (1-\lambda) (\lambda \bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)}) \right] = \\ &= \lambda^2 (1-\lambda)^2 \text{Var} \left[(\lambda \bar{Y}^{(n)} - \bar{Y}_{\lambda}^{(n)}) \right]. \end{aligned}$$

Da mesma forma, utilizando a última proposição, segue que $d_{\lambda}^{(n)*}$ estimador consistente para f_{λ}^* .

Devemos comentar que, para este caso, não conseguimos provar a consistência do estimador proposto, dado o fato de não obtermos resultados análogos ao Teorema de Cramér para variáveis aleatórias não correlacionadas.

4. Estudos de simulação.

Os estudos de simulação neste caso de processo de ramificação ficaram prejudicados. Tivemos problemas para adaptar, em termos de simulação, o fato de não podermos, fixado um número n de observações, garantir a não extinção do processo até a n -ésima geração sem modificar a média do número de descendentes. Na prática, podemos ter extinção do processo em qualquer momento, mas vamos aplicar o estimador no número de observações disponíveis, assumindo que ocorreu uma mudança de média antes da observação n .

O que acontece quando vamos fazer a simulação é que, fixado um tamanho de amostra n , se simplesmente eliminarmos da sequência toda as subsequências que não atingiram o tamanho desejado, estamos de alguma forma selecionando amostras particulares e

isto pode distorcer os resultados, pois não podemos mais garantir as propriedades do gerador de números pseudo-aleatórios.

Por outro lado, se gerarmos um processo de ramificação sem extinção, alteramos nossos estudos, pois estaremos modificando as propriedades do estimador, começando pela sua esperança que é diretamente afetada e isto prejudica a interpretação dos resultados, já que teremos disparidade com a teoria desenvolvida.

Considerando estes fatos, achamos por bem não inserirmos no trabalho esta parte de simulação para este caso, deixando estes estudos para um aprofundamento maior do problema. Talvez uma solução seja deixar o tempo de extinção como variável aleatória, estudar sua distribuição para podermos deixar compatíveis a simulação com a situação prática.

APÊNDICE A : Tabelas.

Tabela 1. Caso Normal ($\sigma_o^2 = \sigma_e^2 = 1$) - $\theta = 0.2$
Monte Carlo - 2000 simulações.

θ^μ	μ_θ	n	s0	1	med	percentis		emq	I.C. $E(\hat{\theta}_1 - \theta_1)^2$	
						5%	95%			
1.0	1.3	30	.88809182	1	0.47	0.07	0.97	.1883	.1787	.1980
				2	0.47	0.17	0.83	.1215	.1154	.1275
				3	0.47	0.13	0.87	.1215	.1154	.1176
		50	.38945800	1	0.40	0.06	0.98	.1891	.1743	.1939
				2	0.44	0.14	0.84	.1147	.1088	.1207
				3	0.46	0.18	0.82	.1142	.1085	.1199
		100	.01619547	1	0.34	0.03	0.98	.1805	.1705	.1905
				2	0.42	0.15	0.85	.1204	.1140	.1269
				3	0.44	0.19	0.81	.1100	.1044	.1155
	1.5	30	.49106352	1	0.40	0.07	0.97	.1676	.1584	.1768
				2	0.43	0.17	0.83	.1108	.1049	.1168
				3	0.43	0.13	0.87	.1123	.1061	.1184
		50	.41402585	1	0.28	0.06	0.96	.1491	.1397	.1585
				2	0.38	0.16	0.82	.0918	.0863	.0974
				3	0.42	0.16	0.80	.0959	.0905	.1013
		100	.30928761	1	0.22	0.04	0.96	.1010	.0928	.1093
				2	0.31	0.16	0.76	.0655	.0608	.0702
				3	0.37	0.18	0.73	.0688	.0647	.0730
	1.7	30	.48136939	1	0.30	0.10	0.97	.1326	.1239	.1413
				2	0.37	0.17	0.80	.0891	.0837	.0946
				3	0.40	0.13	0.80	.0918	.0863	.0972
		50	.75701894	1	0.22	0.06	0.96	.1021	.0939	.1103
				2	0.30	0.16	0.76	.0656	.0609	.0702
				3	0.36	0.16	0.96	.0744	.0696	.0792
		100	.74017132	1	0.20	0.07	0.90	.0541	.0477	.0604
				2	0.26	0.17	0.65	.0366	.0334	.0399
				3	0.34	0.19	0.65	.0469	.0437	.0501

θ^μ - média pré.

μ_θ - média pós.

s0 - semente de geração.

emq - erro médio quadrático.

med - mediana.

Tabela 2. Caso Normal ($\sigma_0^2 = \sigma_0^2 = 1$) - $\theta = 0.5$

Monte Carlo - 2000 simulações.

θ^μ	μ_θ	n	s0	i	med	percentis		emq	I.C. $E(\hat{\theta}_i - \theta_i)^2$	
						5%	95%			
1.0	1.3	30	.13992343	1	0.53	0.07	0.97	.0894	.0861	.0927
				2	0.50	0.17	0.83	.0396	.0375	.0416
				3	0.50	0.13	0.83	.0408	.0385	.0431
		50	.14184441	1	0.52	0.04	0.98	.0954	.0918	.0991
				2	0.50	0.18	0.84	.0378	.0358	.0399
				3	0.50	0.18	0.82	.0336	.0314	.0357
		100	.53073441	1	0.50	0.04	0.98	.0847	.0809	.0884
				2	0.50	0.20	0.77	.0284	.0266	.0303
				3	0.50	0.26	0.73	.0192	.0178	.0206
	1.5	30	.59304573	1	0.50	0.10	0.97	.0714	.0682	.0747
				2	0.50	0.20	0.80	.0313	.0294	.0332
				3	0.50	0.17	0.83	.0333	.0311	.0354
		50	.62381122	1	0.50	0.06	0.96	.0644	.0610	.0678
				2	0.50	0.24	0.78	.0219	.0203	.0234
				3	0.50	0.26	0.74	.0192	.0183	.0216
		100	.42132518	1	0.50	0.08	0.93	.0455	.0424	.0486
				2	0.50	0.29	0.69	.0126	.0115	.0136
				3	0.50	0.33	0.65	.0092	.0083	.0101
	1.7	30	.58778894	1	0.50	0.10	0.93	.0534	.0504	.0564
				2	0.50	0.27	0.77	.0207	.0192	.0223
				3	0.50	0.23	0.77	.0221	.0203	.0239
		50	.25385631	1	0.50	0.10	0.92	.0448	.0418	.0478
				2	0.50	0.30	0.76	.0144	.0132	.0157
				3	0.50	0.34	0.68	.0124	.0112	.0137
		100	.76749958	1	0.50	0.27	0.74	.0188	.0169	.0208
				2	0.50	0.38	0.62	.0056	.0049	.0062
				3	0.50	0.39	0.60	.0040	.0036	.0044

θ^μ - média pré.

μ_θ - média pós.

s0 - semente de geração.

emq - erro médio quadrático.

med - mediana.

Tabela 2. Caso Binomial (1,p) - $\theta = 0.2$

Monte Carlo - 2000 simulações.

θ^H	μ_θ	n	s0	i	med	percentis		emq	I.C. $E(\hat{\theta}_i - \theta_i)^2$	
						5%	95%			
0.2	0.5	30	.56507152	1	0.37	0.20	0.83	.1051	.0991	.1111
				2	0.40	0.17	0.80	.0972	.0917	.1028
				3	0.30	0.10	0.90	.1199	.1120	.1278
		50	.42893034	1	0.26	0.14	0.84	.0793	.0732	.0855
				2	0.34	0.16	0.76	.0745	.0696	.0793
				3	0.26	0.08	0.92	.0943	.0871	.1015
		100	.81602912	1	0.21	0.12	0.80	.0416	.0366	.0466
				2	0.27	0.17	0.66	.0446	.0410	.0483
				3	0.25	0.08	0.69	.0488	.0441	.0535
0.7	0.9	30	.88606779	1	0.20	0.07	0.87	.0708	.0648	.0769
				2	0.33	0.10	0.83	.0863	.0804	.0921
				3	0.50	0.07	0.83	.1421	.1360	.1482
		50	.07706793	1	0.20	0.06	0.88	.0670	.0609	.0732
				2	0.28	0.12	0.78	.0704	.0651	.0756
				3	0.50	0.04	0.86	.1428	.1363	.1492
		100	.33764371	1	0.20	0.06	0.71	.0443	.0395	.0491
				2	0.24	0.13	0.69	.0416	.0379	.0452
				3	0.44	0.09	0.88	.1401	.1334	.1469
0.5	0.9	30	.56944609	1	0.20	0.10	0.60	.0283	.0245	.0321
				2	0.23	0.13	0.70	.0416	.0374	.0458
				3	0.43	0.07	0.83	.1045	.0988	.1102
		50	.62803888	1	0.20	0.10	0.46	.0164	.0135	.0193
				2	0.20	0.14	0.56	.0214	.0186	.0242
				3	0.40	0.04	0.82	.1002	.0946	.1058
		100	.70523165	1	0.20	0.11	0.35	.0080	.0062	.0097
				2	0.20	0.16	0.43	.0101	.0086	.0116
				3	0.38	0.06	0.78	.0791	.0745	.0838

θ^H - média pré.

μ_θ - média pós.

s0 - semente de geração.

emq - erro médio quadrático.

med - mediana.

Tabela 4. Caso Binomial (1,p) - $\theta = 0.5$

Monte Carlo - 2000 simulações.

θ^μ	μ_θ	n	s0	i	med	percentis		emq	I.C. $E(\hat{\theta}_i - \theta_i)^2$	
						5%	95%			
0.2	0.5	30	.23344923	1	0.53	0.23	0.80	.0269	.0252	.0286
				2	0.53	0.30	0.77	.0213	.0198	.0228
				3	0.50	0.13	0.90	.0474	.0447	.0501
		50	.36789869	1	0.50	0.24	0.84	.0273	.0253	.0292
				2	0.52	0.32	0.74	.0149	.0136	.0161
				3	0.50	0.16	0.90	.0324	.0299	.0349
		100	.76477792	1	0.50	0.26	0.77	.0177	.0161	.0194
				2	0.51	0.38	0.66	.0070	.0063	.0078
				3	0.50	0.31	0.62	.0100	.0087	.0113
0.7	0.9	30	.02478777	1	0.40	0.10	0.80	.0493	.0468	.0518
				2	0.47	0.20	0.77	.0299	.0281	.0318
				3	0.50	0.10	0.80	.0425	.0400	.0449
		50	.94983564	1	0.44	0.10	0.76	.0406	.0380	.0431
				2	0.46	0.20	0.72	.0245	.0227	.0262
				3	0.50	0.10	0.84	.0390	.0364	.0415
		100	.19690957	1	0.48	0.13	0.71	.0260	.0238	.0281
				2	0.49	0.28	0.63	.0120	.0108	.0131
				3	0.50	0.36	0.79	.0181	.0164	.0198
0.5	0.9	30	.41022674	1	0.47	0.20	0.60	.0192	.0177	.0207
				2	0.47	0.27	0.60	.0130	.0119	.0143
				3	0.50	0.10	0.73	.0272	.0251	.0294
		50	.62326678	1	0.48	0.26	0.62	.0108	.0096	.0119
				2	0.48	0.34	0.58	.0060	.0053	.0067
				3	0.50	0.38	0.74	.0139	.0124	.0155
		100	.32080898	1	0.50	0.39	0.58	.0043	.0037	.0050
				2	0.49	0.41	0.55	.0022	.0019	.0024
				3	0.50	0.45	0.63	.0039	.0033	.0045

θ^μ - média pré.

μ_θ - média pós.

s0 - semente de geração.

emq - erro médio quadrático.

med - mediana.

Tabela 5. Distribuição pré : Binomial (10,0.1)

Distribuição pós : Binomial (10,0.3)

Monte Carlo com 1000 simulações.

θ	n	s0	i	$\hat{\theta}_i$	I.C. $E(\hat{\theta}_i)$	ema	I.C. $E \hat{\theta}_i - \theta_i $		
0.9	25	.43911076	1	.730	.718	.742	.172	.162	.182
			2	.738	.628	.848	.183	-	-
			3	.740	.629	.851	.184	-	-
	50	.00077201	1	.780	.770	.790	.122	.114	.129
			2	.837	.786	.888	.078	-	-
			3	.823	.771	.875	.086	-	-
	100	.44844578	1	.826	.820	.833	.075	.071	.080
			2	.886	.874	.898	.019	-	-
			3	.882	.863	.901	.029	-	-
0.4	25	.43726780	1	.435	.430	.439	.045	.042	.047
			2	.386	.346	.426	.055	-	-
			3	.421	.373	.469	.061	-	-
	50	.04550454	1	.420	.417	.423	.026	.025	.028
			2	.393	.378	.408	.040	-	-
			3	.409	.389	.429	.040	-	-
	100	.00634953	1	.410	.409	.412	.013	.013	.014
			2	.398	.393	.403	.013	-	-
			3	.407	.400	.414	.017	-	-

s0 - semente da geração.

ema - erro médio absoluto.

Tabela 6. Distribuição pré : Poisson (2)

Distribuição pós : Geométrica (0.5)

Monte Carlo com 1000 simulações.

θ	n	s0	i	$\hat{\theta}_i$	I.C. $E(\hat{\theta}_i)$	ema	I.C. $E \hat{\theta}_i - \theta_i $		
0.8	25	.92967276	1	.596	.583	.610	.219	.208	.230
			2	.681	.577	.785	.175	-	-
			3	.641	.533	.749	.204	-	-
	50	.67581277	1	.646	.636	.658	.167	.158	.176
			2	.748	.692	.804	.109	-	-
			3	.701	.632	.770	.152	-	-
	100	.32508999	1	.604	.685	.703	.117	.110	.124
			2	.784	.761	.807	.058	-	-
			3	.762	.729	.795	.080	-	-
0.4	25	.41257005	1	.461	.451	.470	.112	.106	.118
			2	.473	.390	.556	.153	-	-
			3	.445	.361	.529	.153	-	-
	50	.53062100	1	.438	.430	.446	.082	.077	.087
			2	.435	.387	.483	.108	-	-
			3	.431	.387	.475	.096	-	-
	100	.54239279	1	.420	.414	.427	.054	.051	.058
			2	.416	.398	.434	.049	-	-
			3	.414	.395	.433	.054	-	-

s0 - semente da geração.

ema - erro médio absoluto.

APÊNDICE B. Gráficos.

B1. Gráficos das distribuições de frequências.

Caso independente Bernoulli - 2000 simulações - $\theta = 0.2$.

$$\text{lambdanh} = \hat{\theta}_2 \quad \bullet \quad \text{veross} = \hat{\theta}_1$$

Gráfico B1.1.

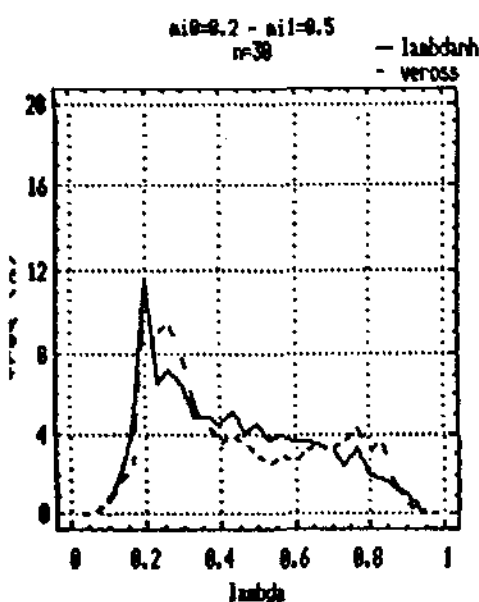


Gráfico B1.2.

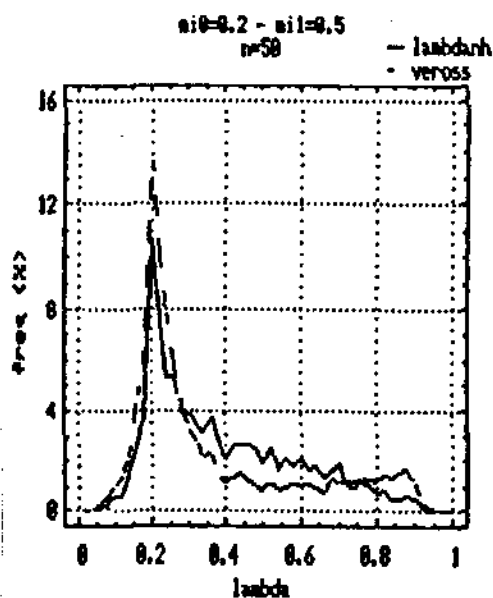


Gráfico B1.3.

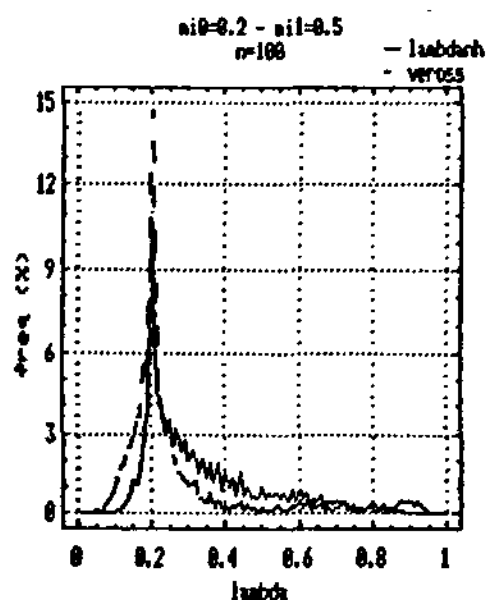


Gráfico B1.4.

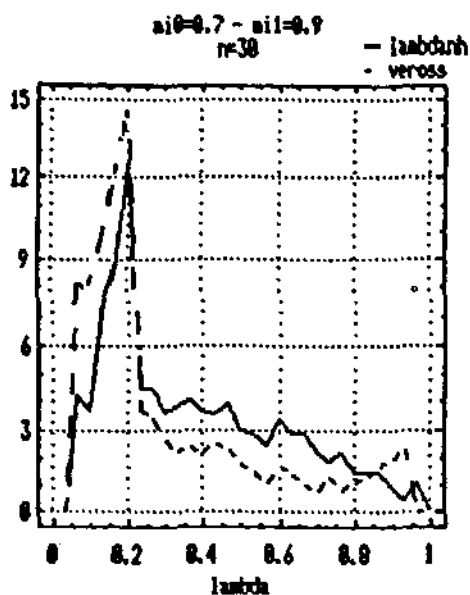


Gráfico B1.5.

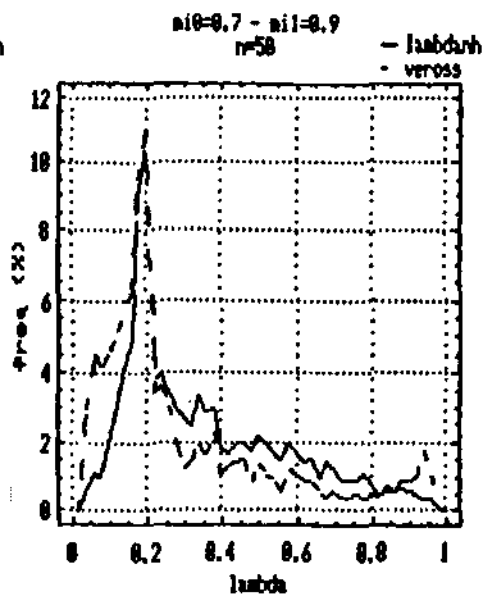


Gráfico B1.6.

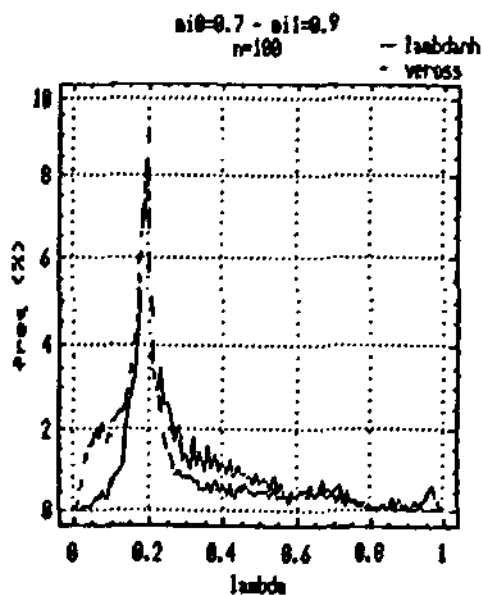


Gráfico B1.7.

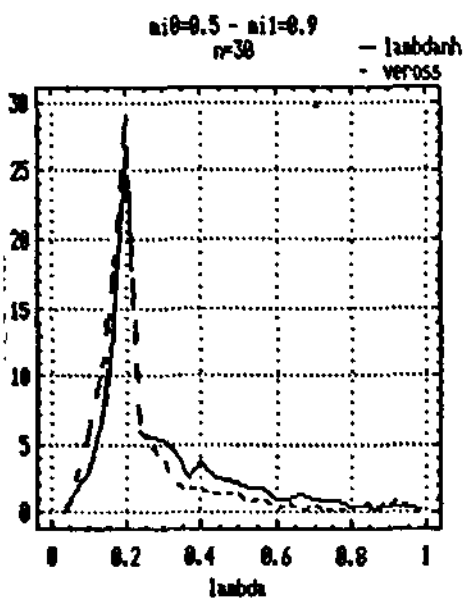


Gráfico B1.8.

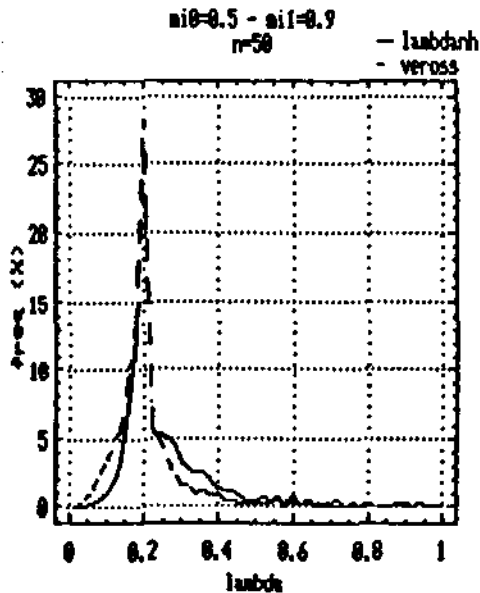
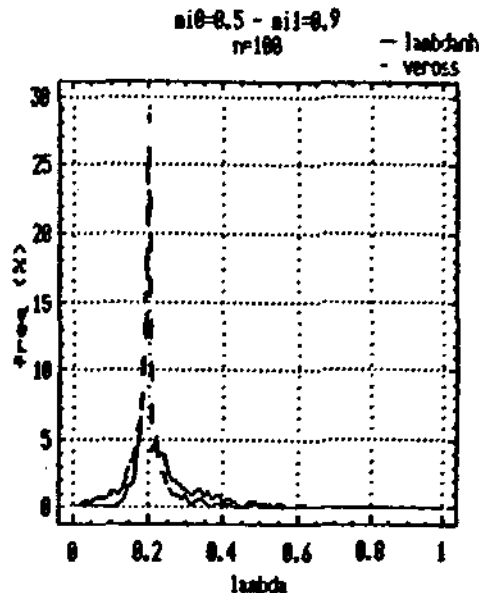


Gráfico B1.9.



B2. Gráficos das distribuições de frequências.

Caso independente Bernoulli - 2000 simulações - $\theta = 0.5$.

$$\text{lambdanh} = \hat{\theta}_2 \quad \text{e} \quad \text{veross} = \hat{\theta}_1.$$

Gráfico B2.1.

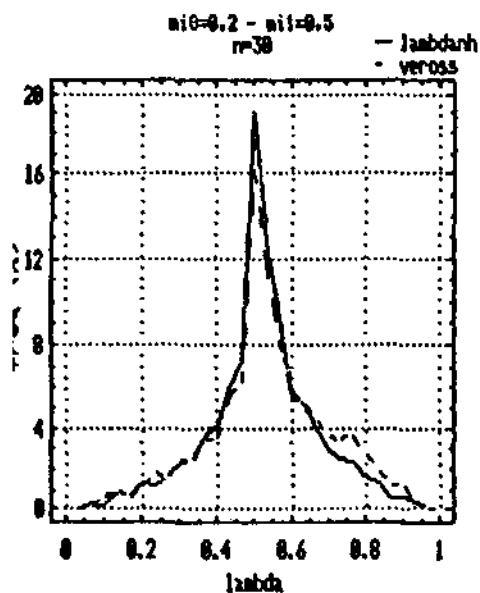


Gráfico B2.2.

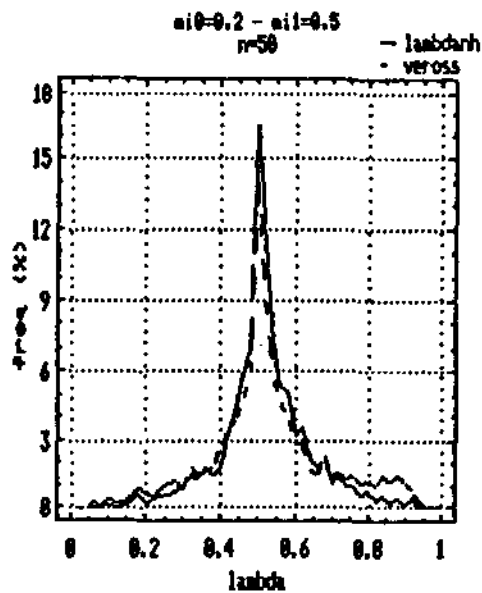


Gráfico B2.3.

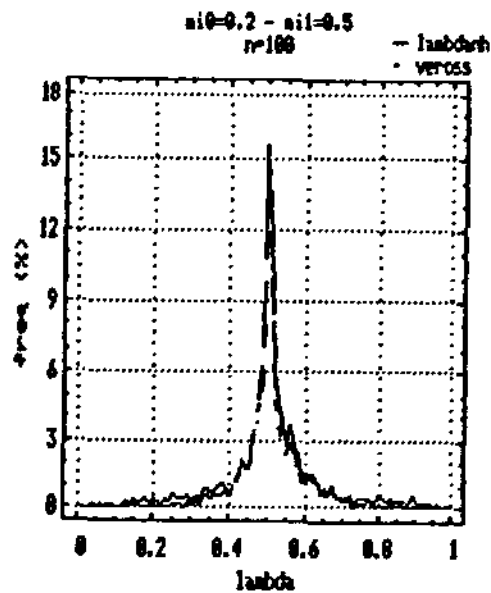


Gráfico B2.4.

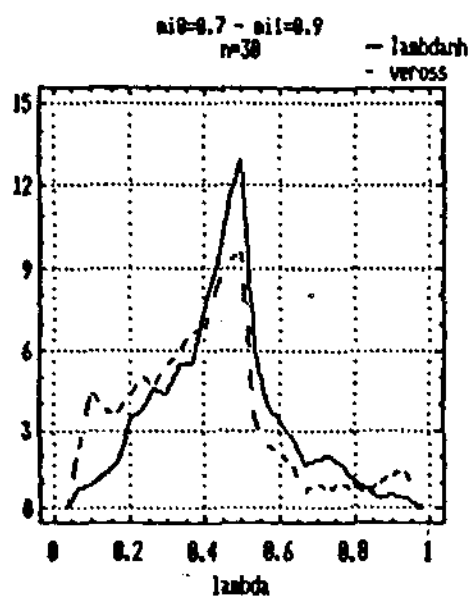


Gráfico B2.5.

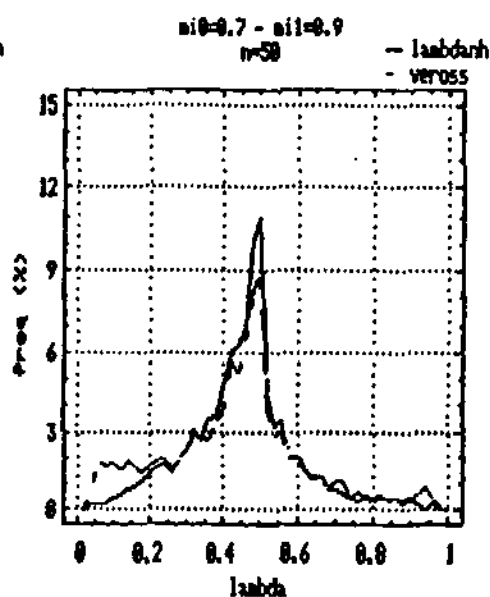


Gráfico B2.6.

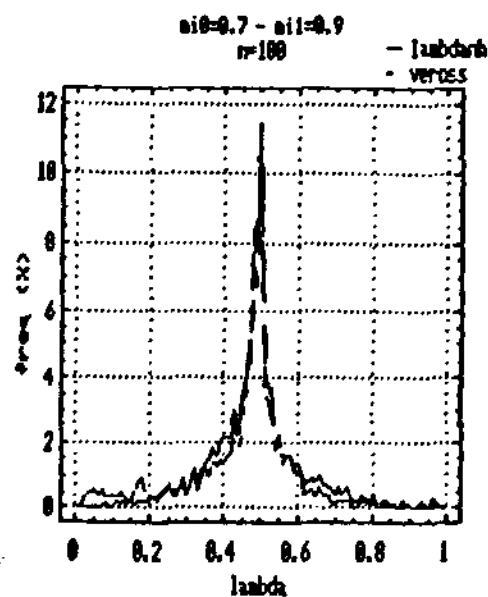


Gráfico B2.7.

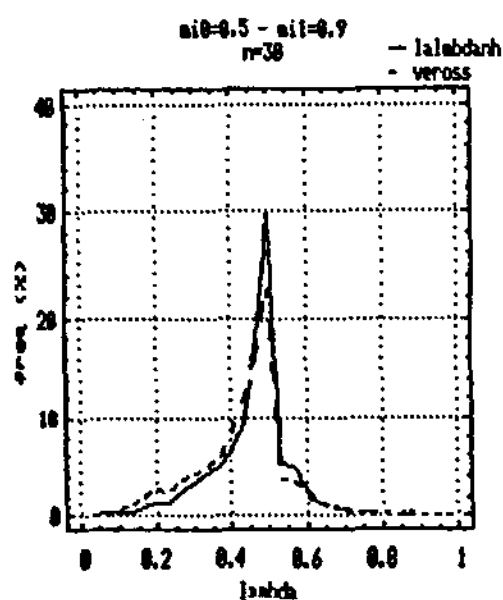


Gráfico B2.8.

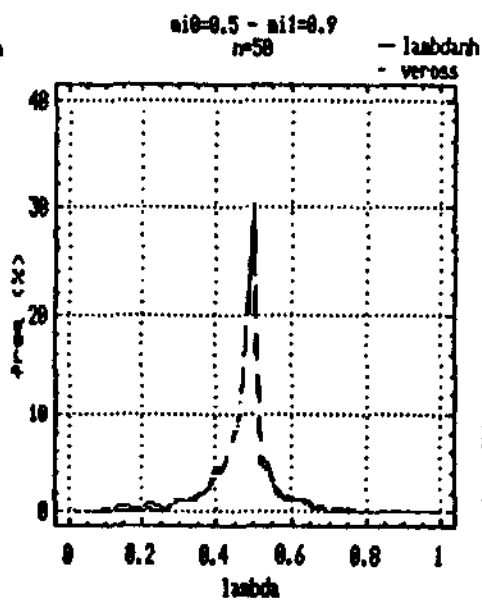
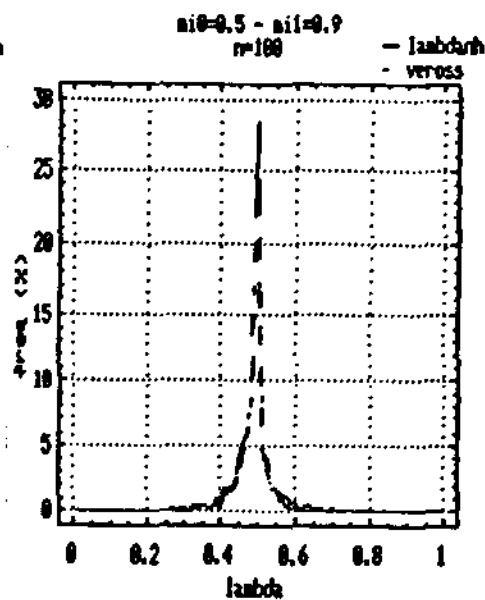


Gráfico B2.9.



B3. Gráficos das distribuições de frequências.

Caso independente Bernoulli - 2000 simulações - $\theta = 0.8$.

$$\text{lambdanh} = \hat{\theta}_2 \quad \text{e} \quad \text{veross} = \hat{\theta}_1.$$

Gráfico B3.1.

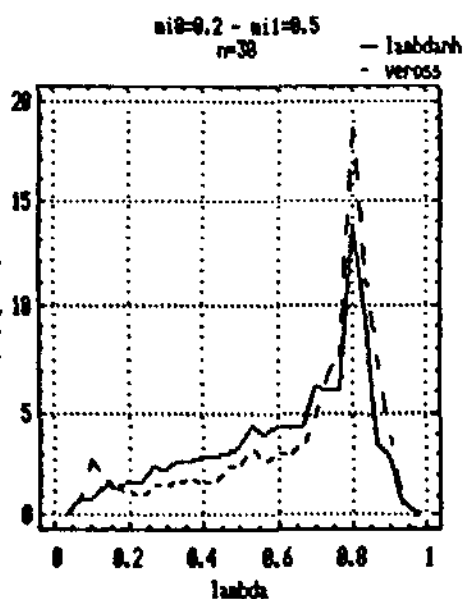


Gráfico B3.2.

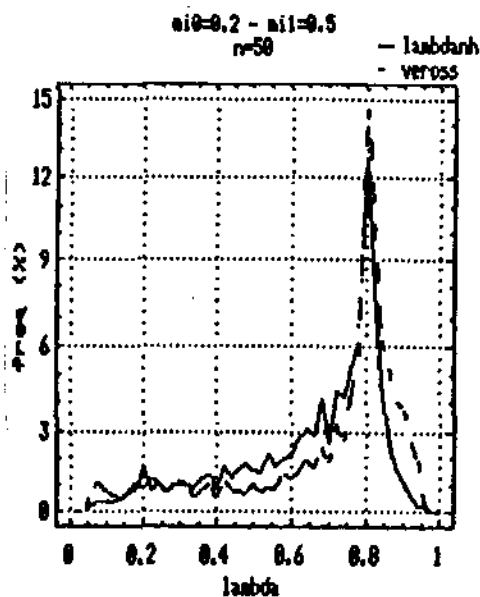


Gráfico B3.3.

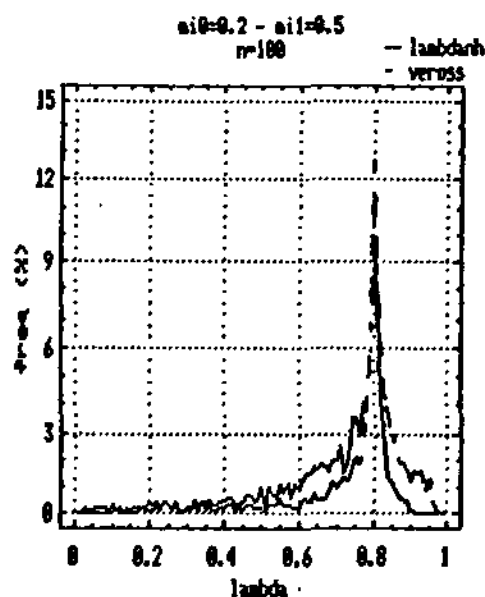


Gráfico B3.4.

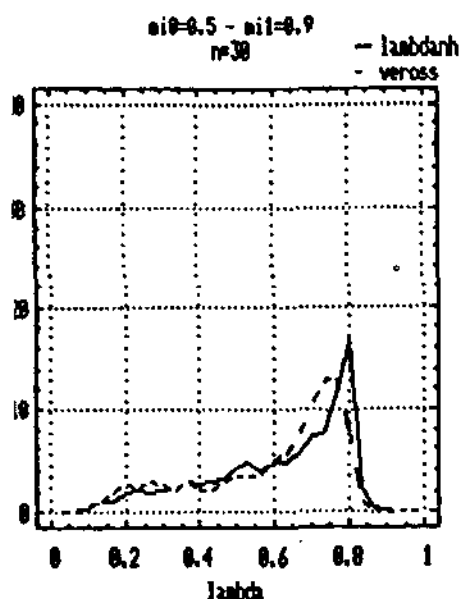


Gráfico B3.5.

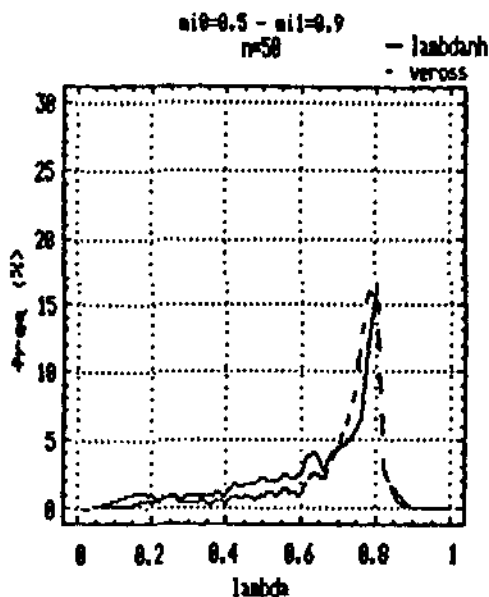


Gráfico B3.6.

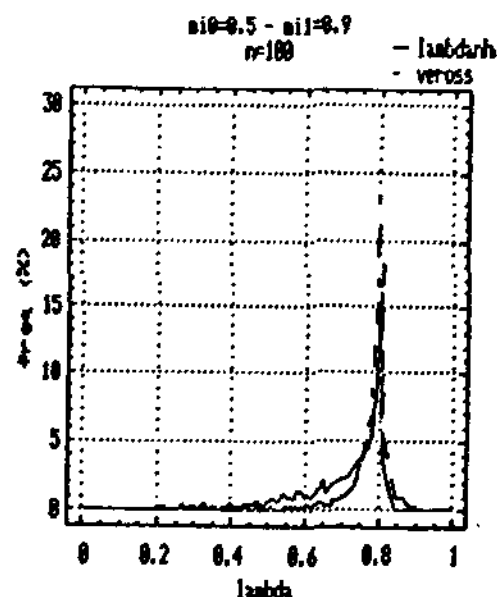


Gráfico B3.7.

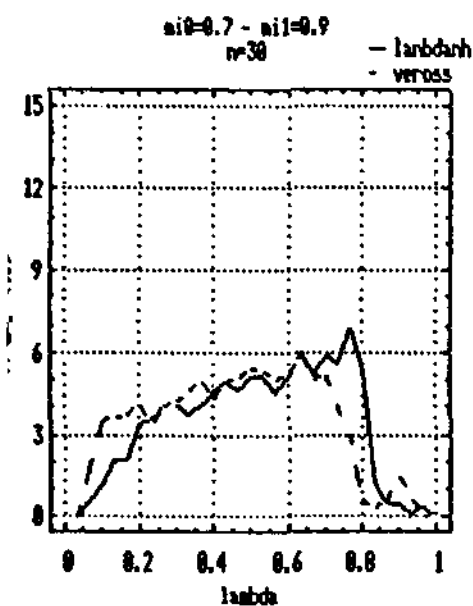


Gráfico B3.8.

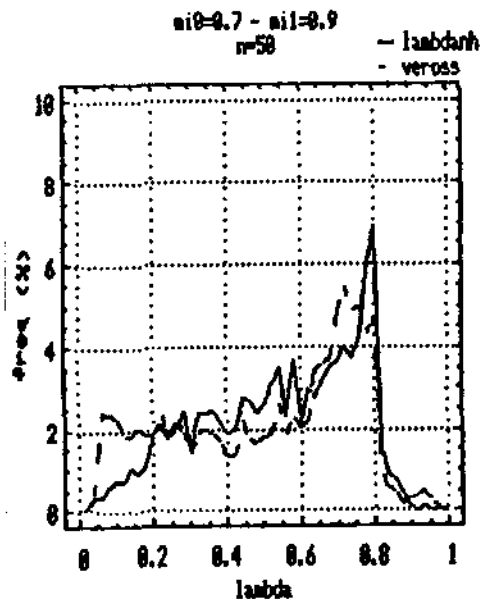
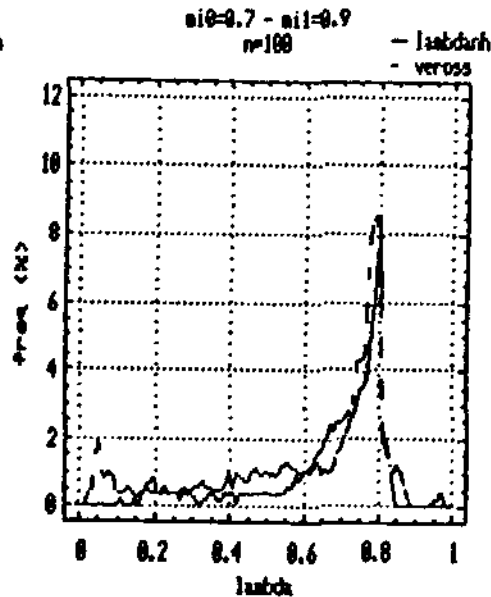


Gráfico B3.9.



APÊNDICE C. Aspectos de simulação.

Cabe aqui ressaltarmos que, para os estudos de simulação feitos neste trabalho, usamos o gerador congruencial linear, definido da seguinte forma (como consta em trabalho feito por Bustos (1991)):

GERADOR CONGRUENCIAL LINEAR.

Seja M inteiro positivo ($M \geq 2$). Uma sequência (y_1, y_2, \dots) em $\{0, 1, \dots, M-1\}$ é dita "gerada pelo gerador congruencial linear com parâmetros a , c e M e semente y_0 " se :

$$y_i = (a y_{i-1} + c) \bmod M, \quad i \geq 1,$$

onde y_0 , a e c são inteiros entre 0 e $M-1$ (incluídos),

a é dito o "multiplicador", c é o "incremento" e

M é o módulo do gerador.

Nós escolhemos um dos geradores ótimos sugeridos por Fishman & Moore (1985), com parâmetros:

$$c = 0,$$

$$a = 950.706.376,$$

$$M = 2^{31} - 1 = 2.147.483.647.$$

APÊNDICE D. Programas.

Programa 1. Caso independente com distribuição bernoulli.

```
PROGRAM artigo1r;
($N+,E+)
  USES CRT;
  LABEL 100;
  CONST      n=100;teta=0.8;nmc=2000;a1=1;a2=1;b1=0.9;b2=0.5;
  VAR        y,nteta,aux,i,j,k:INTEGER;
             u,medpre,medpos,d,dmax,meddmax,varmax,meddes,med1,med2:EXTENDED;
             dh,dhmax,medhmax,varhmax,medhdes,medp1h,medp2h:EXTENDED;
             s2pre,s2pos,dp,dpmax,medpmax,varpmax,medpdes:EXTENDED;
             s,s0,mmed1,mmed2,vardes,varhdes,varpdes,varp1h,varp2h:EXTENDED;
             p1h,p2h:EXTENDED;
             x,freq,freqh,freqp:ARRAY1..100 OF INTEGER;
             imax,ihmax,ipmax:ARRAY1..3000 OF INTEGER;
             arq:TEXT;
```

* FUNÇÃO GERADORA DE NÚMEROS ALEATÓRIOS *

```
FUNCTION RAND(r:EXTENDED):EXTENDED;
VAR a,m:EXTENDED;
BEGIN
  a:=950706376;m:=2147483647;
  s:=INT(FRAC(a*r/m)*m);
  RAND:=s/m;
END;
```

* FUNÇÃO GERADORA DE AMOSTRA COM DISTRIBUIÇÃO BINOMIAL *

```
FUNCTION BIN(a:INTEGER;p:REAL):INTEGER;
BEGIN
  aux:=0;
  FOR j:=1 TO a DO
    BEGIN
      u:=RAND(s);
      IF (u<p) THEN aux:=aux+1;
    END;
  BIN:=aux;
END;
```

* PROGRAMA PRINCIPAL *

```
BEGIN
```

```

CLRSCR;
ASSIGN(arq,'artr3433.DAT');REWRITE(arq);
FOR k:=1 TO nmc DO BEGIN imax[k]:=0;ihmax[k]:=0;ipmax[k]:=0 END;
nteta:=TRUNC(n*teta);
WRITELN(arq,'teta = ',teta:3:1,'      nteta = ',nteta:3,'      n = ',n:4,
',      nmc = ',nmc:5);
WRITELN(arq);
RANDOMIZE;
s0:=RANDOM;
WRITELN(arq,'semente da geracao = ',s0:10:8);
s:=s0;
mmed1:=0;mmed2:=0;medp1h:=0;medp2h:=0;varp1h:=0;varp2h:=0;
FOR k:=1 TO nmc DO
BEGIN
GOTOXY(5,10);WRITE('      ');WRITE('K = ',k:3);
dmax:=0;med1:=0;med2:=0;dhmax:=-1000000;dpmax:=0;

```

*** GERAÇÃO DA AMOSTRA COM PONTO DE MUDANÇA DE DIST. EM NTETA ***

```

FOR i:=1 TO nteta DO BEGIN x[i]:=BIN(a1,b1);med1:=med1+x[i] END;
FOR i:=nteta+1 TO n DO BEGIN x[i]:=BIN(a2,b2);med2:=med2+x[i] END;
mmed1:=mmed1+med1/nteta;
mmed2:=mmed2+med2/(n-nteta);

```

*** CÁLCULO DAS ESTATÍSTICAS DE REFERÊNCIA ***

```

FOR i:=2 TO n-1 DO
BEGIN
medpre:=0;medpos:=0;s2pre:=0;s2pos:=0;
FOR j:=1 TO i DO BEGIN medpre:=medpre+x[j];s2pre:=s2pre+x[j]*
x[j] END;
medpre:=medpre/i;
FOR j:=i+1 TO n DO BEGIN medpos:=medpos+x[j];s2pos:=s2pos+
x[j]*x[j] END;
medpos:=medpos/(n-i);
d:=i/n*(1-i/n)*ABS(medpos-medpre);
dh:=0;
IF ((medpre=0) OR (medpos=0) OR (medpre=1) OR (medpos=1)) THEN
dh:=0
ELSE BEGIN
dh:=i*medpre*LN(medpre)+(i-i*medpre)*LN(1-medpre);
dh:=dh+(n-i)*medpos*LN(medpos)+(n-i-(n-i)*medpos)*
LN(1-medpos);
END;
IF i=n-1 THEN BEGIN dp:=0;GOTO 100 END
ELSE s2pos:=(s2pos-(n-i)*medpos*medpos)/(n-i-1);
s2pre:=(s2pre-i*medpre*medpre)/(i-1);
dp:=d/SQRT(s2pre/i+s2pos/(n-i));

```

* CÁLCULO DOS MÁXIMOS *

```

100: IF dp>dpmax THEN BEGIN dpmax:=dp;ipmaxkl:=1 END;
    IF d>dmax THEN BEGIN dmax:=d;imaxkl:=1 END;
    IF (dh<>0) AND (dh>dhmax) THEN BEGIN
                                dhmax:=dh;
                                ihmaxkl:=1;
                                p1h:=medpre;
                                p2h:=medpos
                                END;

    END;
    medp1h:=medp1h+p1h;
    medp2h:=medp2h+p2h;
    varp1h:=varp1h+SQR(p1h);
    varp2h:=varp2h+SQR(p2h)
END;
medp1h:=medp1h/nmc;
medp2h:=medp2h/nmc;
varp1h:=(varp1h+nmc*medp1h)/(nmc-1)/nmc;
varp2h:=(varp2h+nmc*medp2h)/(nmc-1)/nmc;
FOR i:=1 TO n DO BEGIN freqfil:=0;freqhil:=0;freqpil:=0 END;
medmax:=0;varmax:=0;meddes:=0;
medhmax:=0;varhmax:=0;medhdes:=0;
medpmax:=0;varpmax:=0;medpdes:=0;

```

* DISTRIBUIÇÃO DE FREQUÊNCIAS *

```

FOR k:=1 TO nmc DO
begin
freqimaxkl:=freqimaxkl+1;
freqhihmaxkl:=freqhihmaxkl+1;
freqpilpmaxkl:=freqpilpmaxkl+1;
medmax:=medmax+imaxkl/n;
meddes:=meddes+SQR(imaxkl/n-teta);
medhmax:=medhmax+ihmaxkl/n;
medhdes:=medhdes+SQR(ihmaxkl/n-teta);
medpmax:=medpmax+ipmaxkl/n;
medpdes:=medpdes+SQR(ipmaxkl/n-teta)
end;
medmax:=medmax/nmc;
meddes:=meddes/nmc;
medhmax:=medhmax/nmc;
medhdes:=medhdes/nmc;
medpmax:=medpmax/nmc;
medpdes:=medpdes/nmc;
nmed1:=nmed1/nmc;
nmed2:=nmed2/nmc;
varmax:=0;vardes:=0;varhmax:=0;varhdes:=0;
varpmax:=0;varpdes:=0;

```

```

FOR i:=1 TO n DO
  BEGIN
    varmax:=varmax+freqfil*SQR(1/n-medmax);
    varhmax:=varhmax+freqhfil*SQR(1/n-medhmax);
    varpmax:=varpmax+freqpfil*SQR(1/n-medpmax);
    vardes:=vardes+freqfil*SQR(SQR(1/n-teta)-meddes);
    varhdes:=varhdes+freqhfil*SQR(SQR(1/n-teta)-medhdes);
    varpdes:=varpdes+freqpfil*SQR(SQR(1/n-teta)-medpdes);
  END;
varmax:=varmax/(nmc-1)/nmc;
varhmax:=varhmax/(nmc-1)/nmc;
varpmax:=varpmax/(nmc-1)/nmc;
vardes:=vardes/(nmc-1)/nmc;
varhdes:=varhdes/(nmc-1)/nmc;
varpdes:=varpdes/(nmc-1)/nmc;

```

* IMPRESSÃO *

END.

Principais variáveis utilizadas no programa :

freq : vetor de frequências de $\hat{\theta}_1$.

freqh : vetor de frequências de $\hat{\theta}_2$.

freqp : vetor de frequências de $\hat{\theta}_3$.

medmax : média de $\hat{\theta}_1$ - varmax : variância da média de $\hat{\theta}_1$.

medhmax : média de $\hat{\theta}_2$ - varhmax : variância da média de $\hat{\theta}_2$.

medpmax : média de $\hat{\theta}_3$ - varpmax : variância da média de $\hat{\theta}_3$.

meddes : erro médio quadrático de $\hat{\theta}_1$.

vardes : variância do e.m.q. de $\hat{\theta}_1$.

medhdes : erro médio quadrático de $\hat{\theta}_2$.

varhdes : variância do e.m.q. de $\hat{\theta}_2$.

medpdes : erro médio quadrático de $\hat{\theta}_3$.

varpdes : variância do e.m.q. de $\hat{\theta}_3$.

Programa 2. Caso independente com distribuição normal.

```

PROGRAM artigo2n;
($N+,E+)
    USES CRT;
LABEL 100;
CONST      n=100;teta=0.2;nmc=2000;
VAR        y,nteta,aux,i,j,k:INTEGER;
           u,medpre,medpos,d,dmax,medmax,varmax,meddes,med1,med2:EXTENDED;
           dh,dhmax,medhmax,varhmax,medhdes,medp1h,medp2h:EXTENDED;
           s2pre,s2pos,dp,dpmax,medpmax,varpmax,medpdes:EXTENDED;
           s,s0,mmed1,mmed2,vardes,varhdes,varpdes,varp1h,varp2h:EXTENDED;
           p1h,p2h:EXTENDED;
           freq,freqh,freqp:ARRAY1..1001 OF INTEGER;
           imax,ihmax,ipmax:ARRAY1..30001 OF INTEGER;
           x:ARRAY1..1001 OF EXTENDED;
           arq:TEXT;

    * FUNÇÃO GERADORA DE NÚMEROS ALEATÓRIOS *

FUNCTION RAND(r:EXTENDED):EXTENDED;
VAR a,m:EXTENDED;
BEGIN
    a:=950706376;m:=2147483647;
    s:=INT(FRAC(a*r/m)*m);
    RAND:=s/m;
END;

    * FUNÇÃO GERADORA DE AMOSTRA COM DIST. NORMAL *

FUNCTION NORMAL(b:REAL;c:REAL):EXTENDED;
VAR v1,v2,w,d:EXTENDED;
BEGIN
    v1:=RAND(s);
    v2:=RAND(s);
    d:=SQRT(-2*LN(v1))*sin(6.1432*v2);
    NORMAL:=c*d+b
END;

    * PROGRAMA PRINCIPAL *

BEGIN
    CLRSCR;
    ASSIGN(arq,'artn2133.DAT');REWRITE(arq);
    FOR k:=1 TO nmc DO BEGIN imax[k]:=0;ihmax[k]:=0;ipmax[k]:=0 END;
    nteta:=TRUNC(n*teta);
    WRITELN(arq,'teta = ',teta:3:1,'          nteta = ',nteta:3,'          n = ',n:4,

```

```

      nmc = ',nmc:5);
WRITELN(arq);
RANDOMIZE;
s0:=RANDOM;
WRITELN(arq);
WRITELN(arq,'semente da geracao = ',s0:10:8);
s:=s0;
mmed1:=0;mmed2:=0;medp1h:=0;medp2h:=0;varp1h:=0;varp2h:=0;
FOR k:=1 TO nmc DO
BEGIN
GOTOXY(5,10);WRITE(' ');WRITE('K = ',k:3);
dmax:=0;med1:=0;med2:=0;dhmax:=-1000000;dpmax:=0;

```

*** GERAÇÃO DA AMOSTRA COM PONTO DE MUDANÇA DE DIST. EM NTETA ***

```

FOR i:=1 TO nteta DO BEGIN xil:=NORMAL(1.0,1);med1:=med1+xil END;
FOR i:=nteta+1 TO n DO BEGIN xil:=NORMAL(1.5,1);med2:=med2+xil END;
mmed1:=mmed1+med1/nteta;
mmed2:=mmed2+med2/(n-nteta);

```

*** CÁLCULO DAS ESTATÍSTICAS DE REFERÊNCIA ***

```

FOR i:=2 TO n-1 DO
BEGIN
medpre:=0;medpos:=0;s2pre:=0;s2pos:=0;
FOR j:=1 TO i DO BEGIN medpre:=medpre+x[j];s2pre:=s2pre+x[j]*x[j]
END;
medpre:=medpre/i;
FOR j:=i+1 TO n DO BEGIN medpos:=medpos+x[j];s2pos:=s2pos+x[j]*
*x[j] END;
medpos:=medpos/(n-i);
d:=i/n*(1-i/n)*ABS(medpos-medpre);
dh:=i*(n-i)*SQR(medpre-medpos)/n;
IF i=n-1 THEN BEGIN dp:=0;GOTO 100 END
ELSE s2pos:=(s2pos-(n-i)*medpos*medpos)/(n-i-1);
s2pre:=(s2pre-i*medpre*medpre)/(i-1);
dp:=d/SQRT(s2pre/i+s2pos/(n-i));

```

*** CÁLCULO DOS MÁXIMOS ***

```

100: IF dp>dpmax THEN BEGIN dpmax:=dp;ipmax[k]:=i END;
IF d>dmax THEN BEGIN dmax:=d;imax[k]:=i END;
IF dh>dhmax THEN BEGIN
      dhmax:=dh;
      ihmax[k]:=i;
      p1h:=medpre;
      p2h:=medpos
END;
END;

```



```

medp1h:=medp1h+p1h;
medp2h:=medp2h+p2h;
varp1h:=varp1h+SQR(p1h);
varp2h:=varp2h+SQR(p2h)
END;
medp1h:=medp1h/nmc;
medp2h:=medp2h/nmc;
varp1h:=(varp1h+nmc*medp1h)/(nmc-1)/nmc;
varp2h:=(varp2h+nmc*medp2h)/(nmc-1)/nmc;
FOR i:=1 TO n DO BEGIN freqf1:=0;freqhf1:=0;freqpf1:=0 END;
medmax:=0;varmax:=0;meddes:=0;
medhmax:=0;varhmax:=0;medhdes:=0;
medpmax:=0;varpmax:=0;medpdes:=0;

```

* DISTRIBUIÇÃO DE FREQUÊNCIAS *

```

FOR k:=1 TO nmc DO
  begin
    freqf1max[k]:=freqf1max[k]+1;
    freqhf1max[k]:=freqhf1max[k]+1;
    freqpf1max[k]:=freqpf1max[k]+1;
    medmax:=medmax+1max[k]/n;
    meddes:=meddes+SQR(1max[k]/n-teta);
    medhmax:=medhmax+1hmax[k]/n;
    medhdes:=medhdes+SQR(1hmax[k]/n-teta);
    medpmax:=medpmax+1pmax[k]/n;
    medpdes:=medpdes+SQR(1pmax[k]/n-teta)
  end;
medmax:=medmax/nmc;
meddes:=meddes/nmc;
medhmax:=medhmax/nmc;
medhdes:=medhdes/nmc;
medpmax:=medpmax/nmc;
medpdes:=medpdes/nmc;
nmed1:=nmed1/nmc;
nmed2:=nmed2/nmc;
varmax:=0;vardes:=0;varhmax:=0;varhdes:=0;
varpmax:=0;varpdes:=0;
FOR i:=1 TO n DO
  BEGIN
    varmax:=varmax+freqf1*SQR(1/n-medmax);
    varhmax:=varhmax+freqhf1*SQR(1/n-medhmax);
    varpmax:=varpmax+freqpf1*SQR(1/n-medpmax);
    vardes:=vardes+freqf1*SQR(SQR(1/n-teta)-meddes);
    varhdes:=varhdes+freqhf1*SQR(SQR(1/n-teta)-medhdes);
    varpdes:=varpdes+freqpf1*SQR(SQR(1/n-teta)-medpdes);
  END;
varmax:=varmax/(nmc-1)/nmc;
varhmax:=varhmax/(nmc-1)/nmc;

```

```

varpmax:=varpmax/(nmc-1)/nmc;
varides:=varides/(nmc-1)/nmc;
varhdes:=varhdes/(nmc-1)/nmc;
varpdes:=varpdes/(nmc-1)/nmc;

```

* IMPRESSÃO *

```

CLOSE(arq)
END.

```

As mesmas variáveis do programa 1 foram utilizadas.

Programa 3. Caso independente com distribuições normal, poisson, binomial negativa e normal.

```

PROGRAM arttrg;
($N+,E+)
  USES CRT;
  LABEL 100;
  CONST   n=30;teta=0.9;nmc=1000;
  VAR     y,nteta,aux,i,j,k:INTEGER;
          u,medpre,medpos,d,dmax,medmax,varmax,meddes,med1,med2:EXTENDED;
          s,s0,mmed1,mmed2,varides:EXTENDED;
          x,freq:ARRAY[1..100] OF INTEGER;
          imax:ARRAY[1..3000] OF INTEGER;
          arq:TEXT;

```

* FUNÇÃO GERADORA DE NÚMEROS ALEATÓRIOS *

```

FUNCTION  RAND(r:EXTENDED):EXTENDED;
VAR  a,m:EXTENDED;
BEGIN
  a:=950706376;m:=2147483647;
  s:=INT(FRAC(a*r/m)*m);
  RAND:=s/m;
END;

```

* FUNÇÃO GERADORA DE AMOSTRA COM DIST. BINOMIAL *

```

FUNCTION  BIN(a:INTEGER;p:REAL):INTEGER;
BEGIN
  aux:=0;
  FOR j:=1 TO a DO
    BEGIN
      u:=RAND(s);
      IF (u<p) THEN aux:=aux+1;

```

```

    END;
    BIN:=aux
    END;

```

*** FUNÇÃO GERADORA DE AMOSTRA COM DIST. BINOMIAL NEGATIVA ***

```

FUNCTION      BINN(b:INTEGER;p:EXTENDED):INTEGER;
LABEL 2;
VAR  m:INTEGER;
BEGIN
    aux:=0;m:=0;
2:   u:=RAND(s);
    IF (u>p) THEN BEGIN aux:=aux+1;GOTO 2 END;
    m:=m+1;
    IF m<b THEN GOTO 2;
    BINN:=aux
    END;

```

*** FUNÇÃO GERADORA DE AMOSTRA COM DIST. POISSON ***

```

FUNCTION      POIS(b:REAL):INTEGER;
VAR  nq:INTEGER;
     q,c:REAL;
BEGIN
    q:=1;nq:=0;c:=EXP(-b);
    REPEAT
        u:=RAND(s);
        q:=q*u;nq:=nq+1;
    UNTIL  q<c;
    POIS:=nq-1
    END;

```

*** FUNÇÃO GERADORA DE AMOSTRA COM DISTRIBUIÇÃO NORMAL ***

```

FUNCTION      NORMAL(b:REAL;c:REAL):EXTENDED;
VAR  v1,v2,w,d:EXTENDED;
BEGIN
    v1:=RAND(s);
    v2:=RAND(s);
    d:=SQRT(-2*LN(v1))*sin(6.1432*v2);
    NORMAL:=b*d+c
    END;

```

*** PROGRAMA PRINCIPAL ***

```

BEGIN
    CLRSCR;
    ASSIGN(arq,'artrg111.DAT');REWRITE(arq);
    FOR k:=1 TO nmc DO imax(k):=0;
    nteta:=TRUNC(nsteta);

```

```

WRITELN(arq,'teta = ',teta:3:1,'      nteta = ',nteta:3,'      n = ',n:4,
'      nmc = ',nmc:5);
WRITELN(arq);
RANDOMIZE;
s0:=RANDOM;
WRITELN(arq,'semente da geracao = ',s0:10:8);
s:=s0;
mmed1:=0;mmed2:=0;
FOR k:=1 TO nmc DO
BEGIN
GOTOXY(5,10);WRITE(' ');WRITE('K = ',k:3);
dmax:=0;med1:=0;med2:=0;

```

*** GERAÇÃO DE AMOSTRA COM PONTO DE MUDANÇA DE DIST. EM NTETA ***

```

FOR i:=1 TO nteta DO BEGIN xfi:=BIN(10,0.1);med1:=med1+xfi END;
FOR i:=nteta+1 TO n DO BEGIN xfi:=BIN(10,0.3);med2:=med2+xfi END;
mmed1:=mmed1+med1/nteta;
mmed2:=mmed2+med2/(n-nteta);
FOR i:=2 TO n-1 DO
BEGIN
medpre:=0;medpos:=0;
FOR j:=1 TO i DO medpre:=medpre+xfj;
medpre:=medpre/i;
FOR j:=i+1 TO n DO medpos:=medpos+xfj;
medpos:=medpos/(n-i);
d:=i/n*(1-i/n)*ABS(medpos-medpre);
IF d>dmax THEN BEGIN dmax:=d;imax(k):=i END;
END;
END;
FOR i:=1 TO n DO freqil:=0;
medmax:=0;varmax:=0;meddes:=0;

```

*** DISTRIBUIÇÃO DE FREQUÊNCIAS ***

```

FOR k:=1 TO nmc DO
begin
freqimax(k):=freqimax(k)+1;
medmax:=medmax+imax(k)/n;
meddes:=meddes+ABS(imax(k)/n-teta);
end;
medmax:=medmax/nmc;
meddes:=meddes/nmc;
mmed1:=mmed1/nmc;
mmed2:=mmed2/nmc;
varmax:=0;vardes:=0;
FOR i:=1 TO n DO
BEGIN
varmax:=varmax+freqil*SQR(1/n-medmax);

```

```

    vardes:=vardes+freq[i]*SQR(SQR(1/n-teta)-meddes);
END;
varmax:=varmax/(nmc-1)/nmc;
vardes:=vardes/(nmc-1)/nmc;

```

* IMPRESSÃO *

```

CLOSE(arq)
END.

```

Principais variáveis utilizadas no programa :

freq : vetor de frequências de $\hat{\theta}_1$.
 medmax : média de $\hat{\theta}_1$ - varmax : variância da média de $\hat{\theta}_1$.
 meddes : erro médio absoluto de $\hat{\theta}_1$.
 vardes : variância do e.m.a. de $\hat{\theta}_1$.

OBSERVAÇÕES :

Nos programas utilizados foram calculadas as médias amostrais dos estimadores e respectivas variâncias, além dos erros quadráticos médios e respectivos intervalos de confiança para que pudéssemos analisar os resultados e comparar. Mas depois da simulação feita, observamos que não havia simetria na distribuição de frequências, o que prejudicava demais a análise dos resultados através das médias amostrais. Portanto, tendo em mãos a distribuição de frequência, pudemos calcular as medianas e percentis de 5% e 95%, havendo uma maior coerência com os resultados obtidos.

BIBLIOGRAFIA.

- BAHADUR, R. R. (1971). *Some limit theorems in statistics*. Society for Industrial and Applied Mathematics. Philadelphia.
- BUSTOS, O. (1992). Consulta pessoal.
- CARLSTEIN, E. (1988). Nonparametric change-point estimation. *Ann. Statist.* 16, 188 - 197.
- COBB, G. W. (1978). The problem of the Nile : Conditional solution to a change point problem. *Biometrika* 65, 243 - 251.
- DAGPUNAR, J. (1988). *Principles of Random Variate Generation*. Clarendon Press. Oxford.
- FISHMAN, G. & MOORE, L. (1985). An exhaustive analysis of multiplicative congruential random number generators with modulus $2^{31}-1$. *SIAM Journal of Scientific and Statistical Computing* 7, 24 - 45.
- HINKLEY, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* 57, 1 - 16.
- HINKLEY, D. V. and HINKLEY, E. A. (1970). Inference about the change-point in a sequence of binomial variables. *Biometrika* 57, 477 - 488.
- HOEL, G., PORT, S. G. and STONE, C. J. (1972). *Introduction to Stochastic Processes*. Houghton Mifflin. Boston.
- KARLIN, S. and TAYLOR, H. M. (1975). *A first course in stochastic processes*. Academic Press. London.

- NANTHI, K. (1984).** Statistical estimation for stochastic processes. *Queen' s papers in Pure and Applied Mathematics*. 62.
- PEREZ-ABREU, V. (1989).** Estimación no paramétrica del tiempo de cambio para distribuciones discretas. *Aportaciones Matemáticas* 6, 289 - 301.
- RIPLEY, B. D. (1987).** *Stochastic Simulation*. John Wiley & Sons. New York.
- VARADHAN, S. R. S. (1984).** *Large Deviations and Applications*. Society for Industrial and Applied Mathenmatics. Philadelphia.